

Advancing Ecosystem Monitoring with Global High-Resolution Maps of Vegetation Biophysical Properties

Felix Specker

`felix.specker@usys.ethz.ch`

ETH Zurich <https://orcid.org/0000-0002-9398-9975>

Anna K. Schweiger

Montana State University, Department of Land Resources and Environmental Sciences, Bozeman, MT, United States <https://orcid.org/0000-0002-5567-4200>

Jean-Baptiste Féret

TETIS, INRAE, AgroParisTech, CIRAD, CNRS, Université Montpellier, Montpellier, France
<https://orcid.org/0000-0002-0151-1334>

Luke A. Brown

Department of Geography, King's College London, London, WC2R 2LS, United Kingdom
<https://orcid.org/0000-0003-4807-9056>

Thomas Lauber

Institute of Integrative Biology, Department of Environmental Systems Science, ETH Zurich, Switzerland
<https://orcid.org/0000-0002-3118-432X>

Charbel El Khoury

Swiss Federal Research Institute WSL, Birmensdorf, 8903 Switzerland

Jadunandan Dash

School of Geography and Environmental Science, University of Southampton, Southampton, SO17 1BJ, United Kingdom <https://orcid.org/0000-0002-5444-2109>

Rémi Grousset

ACRI-ST, F-06904, Sophia-Antipolis, France

Bert Gielen

Plants and Ecosystems (PLECO), Department of Biology, University of Antwerp, B-2610, Wilrijk, Belgium
<https://orcid.org/0000-0002-4890-3060>

Thomas W. Crowther

Institute of Integrative Biology, Department of Environmental Systems Science, ETH Zurich, Switzerland
<https://orcid.org/0000-0001-5674-8913>

Johan van den Hoogen

Research Article

Keywords: Ecosystem Monitoring, Biophysical Vegetation Properties, Biodiversity, Restoration, LAI, FAPAR, FCOVER, PROSAIL, Sentinel-2

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-6343364/v3>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Advancing Ecosystem Monitoring with Global High-Resolution Maps of Vegetation Biophysical Properties

Felix Specker^{1,2*}, Anna K. Schweiger³, Jean-Baptiste Féret⁴, Luke A. Brown^{5,6}, Thomas Lauber¹, Charbel El Khoury^{1,2}, Jadunandan Dash⁷, Rémi Grousset⁸, Bert Gielen⁹, Thomas W. Crowther^{1, 10, 11}, Johan van den Hoogen^{1, **}

1. Institute of Integrative Biology, Department of Environmental Systems Science, ETH Zürich, 8092 Zürich, Switzerland

2. Swiss Federal Research Institute WSL, Birmensdorf, 8903 Switzerland

3. Montana State University, Department of Land Resources and Environmental Sciences, Bozeman, MT, United States

4. TETIS, INRAE, AgroParisTech, CIRAD, CNRS, Université Montpellier, Montpellier, France

5. Department of Geography, King's College London, London, WC2R 2LS, United Kingdom

6. School of Science, Engineering & Environment, University of Salford, Manchester, M5 4WT, United Kingdom

7. School of Geography and Environmental Science, University of Southampton, Southampton, SO17 1BJ, United Kingdom

8. ACRI-ST, F-06904, Sophia-Antipolis, France

9. Plants and Ecosystems (PLECO), Department of Biology, University of Antwerp, B-2610 Wilrijk, Belgium

10. King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

11. BRANCH Institute, Zug, Switzerland.

* Correspondence: (felix.specker@wsl.ch)

** Correspondence: (johan.vandenhoogen@usys.ethz.ch)

Summary

Monitoring ecosystem condition and change requires globally consistent, high-resolution vegetation biophysical data. Here, we present S2BIOPHYS, a global 20 m dataset of annual effective leaf area index (LAI_e), fraction of absorbed photosynthetically active radiation (FAPAR), and fractional vegetation cover (FCOVER) for 2019–2025. S2BIOPHYS extends hybrid radiative transfer model inversion by using in-situ reference measurements to jointly optimize the design of training data simulations and neural-network retrieval configurations. Annual products represent near-peak greenness conditions derived from Sentinel-2 observations and include calibrated uncertainty estimates and observation counts, while an accompanying Python package supports retrievals over user-defined periods. Evaluation against >11,000 reference measurements, using ecoregion-blocked cross-validation, shows performance comparable to or exceeding that of benchmark methods within the validation domain. By providing analysis-ready global biophysical variables at field-relevant spatial resolution, S2BIOPHYS can support biodiversity assessment, ecosystem restoration monitoring, and environmental reporting from local to global scales.

Keywords

Ecosystem monitoring; Sentinel-2 imagery; Vegetation biophysical properties; Radiative transfer modelling; Bayesian optimization; Biodiversity

Introduction

Human activities have substantially degraded Earth's ecosystems in recent decades, threatening biodiversity, ecosystem services and human well-being^{1–3}. Conservation, sustainable natural resource management and ecological restoration can partially mitigate these impacts by enhancing CO₂ storage⁴, improving water quality^{5,6} and protecting biodiversity⁷. These efforts gained global momentum with the declaration of the United Nations' 2021–2030 Decade on Ecosystem Restoration, which aims to prevent, halt and reverse ecosystem degradation worldwide⁸. However, monitoring ecological change at relevant spatial and temporal scales remains a major challenge and is essential for evaluating long-term outcomes and directing resources toward effective restoration strategies⁹.

Whereas in-situ monitoring is indispensable for restoration assessments at local scales, remote sensing provides a consistent, scalable means of assessing sustainable or regenerative land management efforts at regional to global scales. Sentinel-2 (S2) data is especially well suited for this purpose, with 13 spectral bands in the visible, near-infrared and shortwave infrared ranges, band-specific spatial resolution of 10 m, 20 m, or 60 m and a revisit time between one (toward the poles) and five days (along the equator). S2 data has proven useful in various ecological applications, including land cover mapping^{10,11}, canopy height estimation through data fusion with LiDAR¹² and retrieval of leaf and canopy traits¹³. Yet, despite the availability of multispectral S2 data, we still lack globally consistent, analysis-ready and high-resolution information on key vegetation biophysical properties that are essential for assessing ecosystem function, degradation, or restoration success.

Radiative transfer models (RTMs) provide a physically based approach for retrieving vegetation biophysical properties from reflectance. By simulating the transfer of electromagnetic radiation through vegetation canopies, RTMs link observed reflectance to leaf- and canopy-level traits, observation geometry and background reflectance¹⁴. A widely used RTM, PROSAIL, couples the PROSPECT leaf model with the SAIL canopy model¹⁵ to simulate canopy reflectance, the fraction of absorbed photosynthetically active radiation (FAPAR) and the fractional vegetation cover (FCOVER), both expressed as fractions ranging from 0 to 1. These outputs depend on four factors: (i) canopy structure, including leaf area index (LAI [m^2m^{-2}], defined as the one-sided green leaf area per unit ground surface area), (ii) leaf optical properties governed by leaf structure and biochemistry, (iii) soil reflectance, and (iv) observation geometry¹⁵. In hybrid inversion approaches, machine learning models are trained on RTM-simulated reflectance to retrieve biophysical variables from satellite observations¹⁴. Although hybrid RTM inversion methods using S2 data have been demonstrated at regional scales^{16–20} and evaluated in global validation studies^{21–23}, these efforts have not yet resulted in globally consistent, precomputed biophysical variable maps at the native spatial resolution of S2.

To fill this gap, we developed a hybrid, open-access retrieval framework that combines radiative transfer modeling with empirically constrained, data-driven optimization informed by a comprehensive in-situ reference dataset. We deployed this framework at the global scale to generate annual maps of three biophysical variables, effective LAI (LAI_e), FAPAR and FCOVER, with calibrated uncertainty estimates at 20 m and 100 m spatial resolutions. These variables are included in the Essential Biodiversity Variables (EBVs)²⁴ framework, with LAI and

FAPAR also recognized as Essential Climate Variables (ECVs)^{25,26}. LAI quantifies the total one-sided green leaf area per unit ground surface and is closely linked to photosynthesis, respiration and evapotranspiration²⁷, while LAI_e represents an effective LAI that ignores canopy clumping while preserving an equivalent spectral response²⁸; FCOVER describes the proportion of a pixel covered by vegetation and is widely used to assess degradation and recovery processes^{29,30}; and FAPAR represents the fraction of incoming photosynthetically active radiation (400-700 nm) absorbed by vegetation and strongly correlates with productivity^{25,31}. Providing these precomputed maps, alongside the Python package *gee-biophys* for retrievals at user-defined temporal frequencies, facilitates downstream applications in biodiversity assessment and improves monitoring of ecosystem restoration worldwide.

Results

Methods Summary

We developed a hybrid RTM retrieval framework in which Bayesian optimization selected the simulated training-data design and feature-engineering choices, while a nested grid search optimized the multi-layer perceptron (MLP) hyperparameters for retrieving LAI, FAPAR, and FCOVER (Figure 1A). The parameter search space (Table S3) included alternative PROSAIL leaf and canopy trait distributions, number of simulated spectra, LAI distributions, added noise, and diverse soil backgrounds derived from the Open-Soil Spectral database³² (OSSL) and from the EMIT imaging spectrometer³³ (NASA's Earth Surface Mineral Dust Source Investigation). For each candidate configuration, we simulated reflectances with the PROSAIL RTM, augmented the training data with spectra from non-vegetated surfaces, and fitted an MLP to retrieve LAI, FAPAR, or FCOVER, respectively.

Model optimization and validation (Figure 1B) used GROUNDED-EO in-situ reference measurements (RMs), compiled from the National Ecological Observatory Network (NEON), Integrated Carbon Observation System (ICOS), and Terrestrial Ecosystem Research Network (TERN). We paired these RMs with spatiotemporally matched, cloud-free S2 observations, retaining a dataset with 11,038 RMs from 1,385 sampling plots across 71 sites, 44 different ecoregions³⁴, and nine biomes³⁵. We used five-fold ecoregion-blocked cross-validation for model training, assigning entire ecoregions to folds, and repeated the optimization described above independently for each split, retaining one best model per split. These five models formed the S2BIOPHYS ensemble. We assessed performance using both stacked out-of-sample retrievals from the held-out folds and ensemble-mean retrievals representing operational model use and compared S2BIOPHYS with two benchmark methods.

To generate annual maps (Figure 1C), we used S2 surface reflectance (L2A) observations³⁶ and applied pixel-level cloud masking using *CloudScorePlus*³⁷. Within each tile and year, we selected S2 acquisitions representing near-peak greenness conditions based on image-level mean Normalized Difference Vegetation Index (NDVI)³⁸ calculated over predominantly natural vegetation, excluding croplands. For applications requiring custom temporal specifications, the *gee-biophys* Python package enables retrievals over user-defined spatial extents and temporal intervals. We aggregated ensemble retrievals into annual pixel-wise means for 2019–2025 at 20 m and 100 m resolution. Within-acquisition uncertainty was estimated from ensemble spread

125 and recalibrated against empirical retrieval errors, while across-acquisition variability was
126 calculated among contributing observations. We combined their variances, assuming zero
127 covariance, to obtain total uncertainty.

128 *Resource Description*

129 The output products consist of annual near-peak greenness composites of LAIe, FAPAR and
130 FCOVER for 2019–2025, with planned annual updates. The datasets are provided at 20 m and
131 100 m resolutions and are hosted on Google Earth Engine (GEE) in S2's native UTM projection
132 using the MGRS tiling grid³⁹. Values are stored as scaled integers using factors of 1,000 for
133 LAIe and 10,000 for FAPAR and FCOVER. All products are publicly accessible via:

```
134 - ee.ImageCollection("projects/ee-speckerfelix/assets/open-  
135   earth/{lai, fapar, fcover}_predictions-mlp_{100, 20}m_v03")
```

136 The 20 m products provide the ensemble mean retrieval (`{var}_mean`) and total retrieval
137 uncertainty (`{var}_stdDev`). At 100m, we additionally provide the within-acquisition uncertainty
138 (`{var}_stdDev_within`), across-acquisition variability (`{var}_stdDev_across`) and the
139 number of valid S2 observations (`{var}_count`), enabling user-defined quality masking.

140 To improve accessibility beyond GEE, we provide global mosaics of the ensemble mean and
141 total uncertainty as Cloud-Optimized GeoTIFFs (COGs) in geographic coordinates
142 (EPSG:4326) at two spatial resolutions (1 km and 100 m)⁴⁰.

143 Each annual image represents the mean across selected near-peak greenness acquisitions and
144 should not be interpreted as an annual mean of vegetation function. Therefore, the annual
145 products should not be used to characterize crop phenology or management-related within-year
146 dynamics. For such applications, the *gee-biophys* Python package ([https://pypi.org/project/gee-
147 biophys/](https://pypi.org/project/gee-biophys/)) should be used. It implements both S2BIOPHYS and SL2P (Simplified Level 2
148 Product Prototype Processor)^{13,19} and lets users define time intervals and spatial extents using a
149 simple configuration file, supporting reproducible and scalable generation of biophysical maps.
150 The package also provides access to the full set of uncertainty components.

151 *Technical validation*

152 We assessed the performance of S2BIOPHYS against in-situ data from GROUNDED-EO
153 RMs⁴¹. Across the represented validation domain (Figure S1), the ensemble-averaged retrievals
154 demonstrated strong performance across all biophysical variables with coefficients of
155 determination (R^2) of 0.77 for LAI, 0.78 for FAPAR and 0.79 for FCOVER and root mean
156 squared errors (RMSE) of 0.60, 0.15 and 0.15, respectively (Table 1, Figure 3). Performance
157 estimates based on stacked out-of-sample retrievals (see methods) was comparable (LAI:
158 $R^2 = 0.75$, RMSE = 0.63; FAPAR: $R^2 = 0.75$, RMSE = 0.16; FCOVER: $R^2 = 0.77$, RMSE =
159 0.15; Table S4). Across variables, retrievals tended to overestimate low values and
160 underestimate high values (Figure 3). Mean errors were small and positive (LAI: +0.07;
161 FAPAR: +0.02; FCOVER: +0.01). The Uncertainty Agreement Ratio (UAR) showed the highest
162 consistency for LAI (90%), followed by FAPAR (70%) and FCOVER (68%; Table S4).

163 For benchmarking, we compared S2BIOPHYS with two retrieval methods applicable to S2
164 observations: SL2P^{13,19}, an established PROSAIL-based inversion model, and the
165 GROUNDED-EO Gaussian Process Regression (GPR) model⁴¹, a data-driven retrieval model
166 trained directly on RMs. S2BIOPHYS outperformed SL2P, showing higher R^2 values (LAI: 0.77
167 vs. 0.68; FAPAR: 0.78 vs. 0.69; and FCOVER: 0.79 vs. 0.74) and lower RMSE values (0.60 vs.
168 0.70, 0.15 vs. 0.18 and 0.15 vs. 0.16, respectively) (Figure 3, Table 1). Against GROUNDED-EO
169 GPR, S2BIOPHYS achieved comparable FAPAR performance ($R^2 = 0.78$ vs. 0.79; RMSE =
170 0.15 vs. 0.15), with small metric-specific differences (Table S4). Map-to-map comparisons
171 across four representative landscapes showed high rank correlations between S2BIOPHYS and
172 benchmark retrievals, despite systematic offsets relative to SL2P or GROUNDED-EO GPR
173 (Figures S2–S3). Aggregation to 300 m reduced fine-scale spatial detail, and rank correlations
174 with Sentinel-3 CLMS^{42,43} were considerably lower than those among the 20 m retrievals with
175 SL2P or GROUNDED-EO GPR.

176 We evaluated variation in retrieval performance among vegetation types by assessing
177 S2BIOPHYS separately for each land-cover class⁴⁴. Because biophysical property ranges differ
178 among classes, we additionally report RMSE normalized by the class-specific mean (nRMSE).
179 S2BIOPHYS had the lowest absolute RMSE in shrub/scrub for all three variables, but the lowest
180 nRMSE in mixed and deciduous forests (Table S4). Compared with SL2P, S2BIOPHYS had
181 lower RMSE for LAI in most land-cover classes, including cultivated crops (RMSE 0.50 vs.

0.80), evergreen forests (0.69 vs. 0.78), and grassland/herbaceous vegetation (0.45 vs. 0.75). For FAPAR, S2BIOPHYS had similar or lower RMSE than SL2P among all land-cover classes, with the largest differences in mixed forest (0.10 vs. 0.17) and deciduous forest (0.15 vs. 0.19). FAPAR retrieval performance of S2BIOPHYS was similar to GROUNDED-EO GPR overall, with lower RMSE in cultivated crops, deciduous forest and wetlands, but higher RMSE in evergreen forest, pasture/hay and shrub/scrub.

We also examined whether model calibration benefited from stratification by comparing group-specific models calibrated separately for eight biome–land-cover groups with a generic model calibrated across all groups (Note S3, Tables S5). Group-specific calibration did not improve pooled performance: the generic model had lower RMSE and higher R^2 for LAIe, FAPAR, and FCOVER. For LAIe, specific models reduced RMSE in the three largest validation groups but performed substantially worse in some smaller groups, particularly cold-climate and tropical evergreen forests (Table S6).

We evaluated how well the raw ensemble spread corresponded to empirical retrieval errors against GROUNDED-EO RMs. Because ensemble spread generally underestimated empirical errors for intermediate and high retrieval values, we recalibrated it using a smooth scaling function of the predicted value (Figure S4). Recalibration improved correspondence with empirical errors: mean absolute calibration error (MACE) decreased from 0.372 to 0.271 for LAIe, from 0.086 to 0.066 for FAPAR and from 0.086 to 0.066 for FCOVER (Figure S4). S2BIOPHYS also had lower MACE than SL2P, which similarly provides pixel-level uncertainties, both before and after recalibration. These results support the use of the recalibrated ensemble spread as an empirically calibrated proxy for within-acquisition retrieval uncertainty.

Whereas this recalibration characterizes retrieval uncertainty for individual acquisitions, annual composites are also affected by variability among their contributing observations. Analysis of the two uncertainty components revealed distinct spatial patterns (Figure S5). Within-acquisition uncertainty was highest in structurally complex and heterogeneous canopies, including dense forests and mixed land-cover boundaries. Across-acquisition variability was elevated in regions with strong seasonal dynamics, active land management, or variable observation conditions, including atmospheric variability. Because this component can include changes in vegetation state as well as observation-related variation, it reflects the stability of contributing observations to the annual composite. Consequently, the relative contributions of within-acquisition

uncertainty and across-acquisition variability differed among ecosystems and land-use types (Figure S5).

Limitations

The parametrization of the model training workflow (Figure 1) relies on in-situ RMs for calibration, which are geographically biased toward North America and Europe (Figure S1). As a result, empirical validation is necessarily weaker in underrepresented regions. To reduce bias in performance estimates arising from spatial autocorrelation, we train the model ensemble using ecologically meaningful, ecoregion-based spatially blocked CV. While this approach aims to obtain more independent validation samples within the represented ecoregions, it cannot compensate for the absence of RMs in underrepresented ecoregions, where model accuracy remains empirically unconstrained. This limitation motivated the use of RTM simulations to promote physical consistency beyond the calibration domain, while underscoring the need for expanded in-situ measurements in underrepresented ecosystems, particularly across Asia, Africa and South America.

To quantify extrapolation effects, we analyzed retrieval performance and uncertainty as functions of spectral distance from the training domain. Retrieval performance declined with increasing distance from the LUT, with a marked drop beyond ~3.5 (Figure S9). Retrieval uncertainty increased with spectral distance, thereby reflecting reduced retrieval reliability, except in the most extreme extrapolation cases. The mean distance to the PROSAIL LUT was similar for GROUNDED-EO RMs and global S2 samples (1.69 and 1.75, respectively), and only 2.5% and 3.5% exceeded the 3.5 threshold. Thus, the validation and prediction domain showed similar degrees of extrapolation from the training domain. Given the substantial performance decline beyond this threshold, we generated a spatially explicit global map of spectral distance to the training domain at 100 m resolution. Strong extrapolation was limited to a few regions, primarily in steep terrain or areas of intensive agriculture, and these pixels can be identified and masked if desired (Figure S9).

The validation data also includes several sources of uncertainty. GROUNDED-EO RMs include per-measurement uncertainty estimates, but these likely represent lower bounds⁴⁵ and additional error arises from spatial footprint mismatch and temporal offsets between measurements and satellite acquisitions (median offset: 2.27 days). In addition, S2-L2A data lack an operational uncertainty product⁴⁶, meaning that uncertainty due to sensor noise,

atmospheric correction errors and bidirectional reflectance distribution function (BRDF) effects (i.e. reflectance changes caused by varying illumination and viewing geometry) cannot be propagated through the retrieval workflow. Empirical retrieval errors therefore reflect a combination of model uncertainty, RM noise, and unquantified S2 observation uncertainty. Our post-hoc uncertainty recalibration adjusts the within-acquisition ensemble uncertainty (see methods) to match this combined error structure, but it cannot compensate for the absence of explicit S2 uncertainty inputs.

The PROSAIL model operates under several assumptions, most notably that of a turbid medium, i.e., a homogeneous, continuous canopy with randomly distributed leaves. Consequently, our retrieved LAI is an *effective* quantity, which describes the LAI of a canopy with randomly distributed leaves that produce an equivalent spectral response. While previous research has shown that SAIL-derived LAI closely matches LAI²³, the distinction between LAI and LAI limits the direct applicability of our retrievals, for example in structurally complex ecosystems, and would require additional estimates of, or corrections for, canopy clumping to derive LAI.

Map quality directly reflects the quality and availability of input S2 data. Regions with insufficient input data due to frequent cloud cover include, e.g., tropical regions in East Africa and Latin America, and parts of the northern Andes (Figure S6). We observe moderate tiling and orbital artifacts in agriculture-dominated areas such as the Midwestern United States and North India, likely due to a temporal mismatch between peak natural and peak agricultural growing seasons (Figure S6). In these cases, we expect the NDVI signal used for image selection to be confounded by agricultural practices due to the lack of remnant natural vegetation or insufficient masking of agricultural land-use pixels. Tiling artifacts are also apparent in FCOVER maps over tropical closed-canopy regions, potentially partly attributable to limitations of Sen2Cor's atmospheric correction⁴⁷, which is known to underperform in high-aerosol tropical environments⁴⁸. Orbital artifacts are visible in the LAI maps, especially over dense canopies such as the Amazon (Figure S6), which are likely caused by BRDF effects⁴⁹. Although PROSAIL accounts for observation geometry in its simulations, angular information accessible within GEE is restricted to the tile level (110x110 km) rather than at the finer scale (5x5 km) from the original ESA S2 product⁴⁷. Consequently, more detailed angular predictors were not included, and residual BRDF effects may remain.

Discussion

Implications and potential applications

Here, we present S2BIOPHYS, a global dataset of 20 m annual near-peak greenness maps of LAI, FAPAR and FCOVER, together with the *gee-biophys* Python package for generating retrievals over user-defined time intervals. Together, these open-source resources support diverse applications in ecosystem monitoring, land management and restoration assessment across spatial and temporal scales, from local case studies to global comparisons and from short-term disturbance assessments to long-term vegetation dynamics. At 20 m resolution, the products capture fine-scale spatial patterns and within-site variability that coarser products may not resolve, supporting analysis of small management units and heterogeneous land-use systems.

Beyond static representations, the annual time series and interannual difference maps enable the detection and quantification of vegetation disturbance and recovery processes across large spatial extents. As demonstrated across forest harvests in Sweden, annual FCOVER maps can quantify disturbance-related biophysical changes at scales down to sub-hectare areas (Figure 4C, Note S2). More generally, year-to-year difference maps can reveal vegetation responses to natural disturbances such as drought and wildfire, as well as anthropogenic disturbances including deforestation, while longer time series facilitate assessments of recovery trajectories following disturbance or restoration (Figure 4B). When combined with complementary datasets such as canopy height or above-ground biomass, these products allow differential recovery rates to be compared across ecosystem dimensions, reflecting the multidimensional nature of forest recovery⁵⁰. Explicit uncertainty estimates, including separate within-acquisition uncertainty and across-acquisition variability components (Figure S5), further support interpretation of temporal changes by separating disagreement among ensemble members from variability among acquisition-level retrievals.

At similarly fine spatial scales, S2BIOPHYS also enables analyses of habitat heterogeneity, which is linked to biodiversity, with greater structural or functional variability generally associated with higher biodiversity^{51–53}. Remotely sensed heterogeneity has been quantified using unimodal metrics (e.g. variability in spectral indices⁵⁴ or canopy height⁵⁵), or multivariate metrics (e.g. functional richness, divergence, evenness⁵⁶). We show that, for example, across 464 matched plantation-forest pairs distributed throughout Tasmania, plantations had higher

mean LAI but significantly lower spatial heterogeneity than nearby forests (Figure 4A, Note S1). This landscape-scale comparison demonstrates the potential of using high-resolution biophysical maps to quantify vegetation heterogeneity across ecosystems and land-use systems, supporting more integrated ecosystem monitoring. By providing globally consistent, high-resolution biophysical variables, S2BIOPHYS enables broad-scale testing of relationships among vegetation structure, functional diversity and biodiversity patterns.

Comparison with existing resources

Although previous mapping products are available for LAI, FAPAR, or FCOVER, they are typically limited in either spatial resolution (> 30 m pixel size) or spatial extent and temporal continuity. For example, the Canada Centre for Remote Sensing (CCRS) provides operational 20 m monthly summer composites based on land-cover-specific modifications of SL2P (SL2P-CCRS⁵⁷) for Canada, while a near-real-time implementation of SL2P for Europe within the Copernicus Land Monitoring Services (CLMS)⁵⁸ was discontinued in 2025. Large-scale mapping efforts at 30 m resolution using the Landsat record have been demonstrated for the contiguous United States, but these products are either available as on-demand processing services⁵⁹ or have not been updated anymore⁶⁰.

In contrast, existing global operational products are usually coarser in spatial resolution but offer higher temporal frequency. For example, the MODIS LAI product⁶¹ provides 500 m resolution with an eight-day frequency, while LAI, FAPAR and FCOVER products derived from Sentinel-3 are available at 300 m through CLMS^{42,43,62}. However, most applications in ecological management and land surface modeling require spatial resolutions that match the scale of land management units and landscape heterogeneity. For example, in smallholder-dominated regions, land parcels often occupy less than five hectares^{63,64}, equivalent to 125 pixels at 20 m but only a fraction of a single pixel at 300 m resolution, highlighting the limitations of coarse-resolution products for monitoring small-scale restoration and management activities^{65,66}.

Similarly, our analysis of forest harvests across Sweden shows that aggregating FCOVER to coarser resolutions progressively attenuates the quantified biophysical change (Figure 4A, Note S2).

To provide high spatial resolution at a global scale while keeping storage requirements tractable, our precomputed product focuses on annual aggregations rather than high-frequency retrievals. For many ecological and land-management applications, annual estimates of

biophysical variables are sufficient to characterize long-term vegetation condition and restoration progress. To enable consistent global comparisons across regions with differing phenology, we defined the growing season empirically using tile-level NDVI aggregations, selecting observations with minimal cloud cover and near-peak NDVI. We observe that the timing of near-peak greenness varies systematically across latitudinal zones (Figure S7), a pattern not captured by products that rely on annual medians or globally fixed time window¹². The annual products should not be used to characterize crop phenology or management-related within-year changes, because crop development may not coincide with the regional near-peak greenness period used for image selection. These applications require sub-annual retrievals, which can be generated with the package *gee-biophys* or alternative retrieval frameworks such as PyEOGPR⁶⁷ and SL2P-based tools in the LEAF Toolbox^{19,68,69}.

Method comparison

Global-scale, high-resolution retrievals require models that balance physical realism, predictive flexibility across heterogeneous surface types, and operational scalability. Hybrid RTM-based inversion approaches, such as S2BIOPHYS and SL2P, address this balance by coupling physically based forward simulations with non-parametric regression for variable retrieval^{14,70}. PROSAIL has been widely adopted for this purpose because of its computational efficiency⁷⁰ and mechanistic representation of canopy structure, leaf biochemistry, and background reflectance. Building on earlier hybrid PROSAIL-based inversion approaches, our framework systematically optimizes RTM parameter distributions using in-situ observations to improve retrieval accuracy and robustness at the global scale.

Increasing model complexity (e.g., using vegetation-type-specific retrieval models^{57,71}) can reduce systematic bias in structurally heterogeneous systems, but is subject to a bias–variance trade-off⁵⁷. Our biome–land-cover-specific calibration experiment reflected this trade-off: specific models improved LAI retrievals in the three largest validation groups, but performed substantially worse in some other groups, while the generic model performed better across all groups pooled for all three variables (Note S3, Table S6). At global scale, vegetation-specific retrievals additionally require accurate, high-resolution land-cover maps to select the appropriate model. However, such maps are often unavailable at the required spatial or temporal resolution^{11,72,73} or use broad classes that combine vegetation types^{11,74,75}. Moreover, classification uncertainties and errors propagate directly into the retrieval chain when conditioning on misclassified labels^{13,41,76}. Dependence on external land-cover products may

further constrain operational deployment through changes in their availability, consistency, or update frequency. We therefore adopt a single, globally applicable PROSAIL-constrained model to prioritize robustness and operational scalability without dependence on external land-cover products.

A further requirement for global retrieval is the ability to represent the full diversity of Earth's surface conditions. Earlier hybrid retrievals accounted for soil variability, for example by selecting representative soil spectra from soil libraries^{13,77}, but did not explicitly integrate the diversity of non-vegetated surfaces encountered globally. We expand this representational domain by incorporating in-situ (OSSL³²) and remotely-sensed (EMIT³³) soil spectra, as well as non-vegetated surface reflectance from S2. While EMIT spectra have been used in albedo and radiative-budget studies⁷⁸, their direct use as background inputs within a PROSAIL-based vegetation retrieval workflow is, to our knowledge, new. This enriched spectral representation improves retrieval robustness in sparsely vegetated, arid and heterogeneous landscapes where background reflectance strongly influences observed canopy signals and helps support more reliable generalization.

Meeting the computational demands of global 20 m mapping further requires a retrieval algorithm that can be deployed efficiently on cloud platforms. We designed S2BIOPHYS for server-side inference within GEE, which informed our algorithmic choices. MLPs can be trained locally, serialized reliably and executed efficiently in GEE, enabling rapid iteration and reproducible deployment. Although alternative approaches such as GPR or Random Forest (RF) have been implemented in GEE^{59,79–81}, we encountered practical challenges in GEE model deployment, including the need for substantial training-data reduction (GPR) and reproducibility limitations related to model serialization of locally trained RF models for server-side execution. MLPs have demonstrated strong performance in several operational biophysical retrieval toolboxes^{13,19,60} and offer a pragmatic balance between accuracy, scalability and implementation simplicity. Overall, integrating a physically grounded RTM backbone with globally enriched and empirically optimized RTM parametrization, combined with efficient cloud-based deployment, positions S2BIOPHYS as a robust and scalable framework for producing global, analysis-ready vegetation biophysical variables.

Methods

Model Training

We performed model training separately for each variable and CV split. For each split, we integrated a hybrid RTM retrieval workflow with Bayesian optimization to jointly optimize RTM input distributions, LUT modifications, and feature engineering specifications (Figure 1A, Table S3). We defined input ranges for foliar and canopy structural traits and generated trait combinations as input for the PROSAIL RTM⁸² (Figure 1A1). Drawing from previous literature, we replicated parametrizations implemented by the Simplified Level 2 Biophysical Processor (SL2P)¹³ and by Wan et al.²¹, Estévez et al.⁸⁰, and Kovács et al.⁸¹ (Table S2). Rather than fixing the LAI distribution, we treated its mean and standard deviation as trainable parameters (Table S3). Because PROSAIL assumes a homogeneous turbid medium without foliage clumping¹⁵, LAI and effective LAI (LAI_e) are equivalent within the model; we therefore used the input LAI values as the LAI_e targets when training the retrieval model. As an additional parameterization option, we accounted for correlations among leaf traits^{83,84} using foliar trait data from NEON (DP1.10026.001)⁸⁵. We used a Gaussian copula function⁸⁶ to model the covariance structure of the measurements for leaf chlorophyll (CHL), leaf carotenoid (CAR) contents, equivalent water thickness (EWT), and leaf mass per area (LMA). Before fitting, we excluded observations where EWT was above 0.1 cm, CHL above 100 $\mu\text{g cm}^{-2}$, LMA above 0.10 g cm^{-2} , or carotenoid content above 30 $\mu\text{g cm}^{-2}$. We modeled the marginal distributions as normal distributions truncated at zero and generated 100,000 samples that preserved the observed covariance structure (Figure S8).

Next, we defined the illumination and observation geometry using either predefined ranges of observation angles or observation angles sampled from randomly selected S2 images collected during the growing season. To represent the soil background, we combined the standard implementation of *psoil*, defined as the weighted average of a dark and bright reference soil spectra, in the *prosail* R package⁸⁷ with two additional datasets: reflectance data from EMIT^{88,89}, which provides spectra at 60 m spatial and 7.5 nm spectral resolution covering the visible to near-infrared (VisNIR, 380 – 2500 nm) region of the electromagnetic spectrum, and VisNIR spectra from OSSL^{32,90–102}, which contains harmonized in-situ soil spectra collected from various instruments. We parametrized the fraction and type of background reflectance used as trainable parameters (Table S3).

We applied the PROSAIL RTM to simulate canopy reflectances (Figure 1A2) and converted the output's 1-nm spectral sampling to the corresponding S2 bands using the S2 spectral response functions¹⁰³. The PROSAIL RTM combines the leaf optical model PROSPECT¹⁰⁴ with the canopy structural model 4SAIL¹⁰⁵. We used the implementation from the *prosail* R package⁸⁷, version 2.4.1, which invokes the PROSPECT-D model¹⁰⁶ in the *prospect* R package¹⁰⁷ version 1.7.0. We included the number of simulated spectra as a trainable parameter (Table S3). This formed a look-up table (LUT) containing input traits as defined in PROSPECT-D (including LAI), output reflectance, and canopy properties (including FAPAR, FCOVER).

We augmented the LUT with reflectance spectra of non-vegetated surfaces to better represent global surface heterogeneity. We incorporated bare soil and rock, snow, ice and built-up spectra from S2 and soil spectral data from EMIT and OSSL. We assigned zero values to LAI, FAPAR, or FCOVER for these non-vegetated spectra and optimized the number of spectra added from each group and the levels of additive and multiplicative noise during training (Table S3). Additional preprocessing options optimized during training included normalizing reflectance using the near-infrared reflectance of vegetation^{21,108} (NIRv), and transforming response variables (LAI, FAPAR, FCOVER) through standard or logarithmic scaling, or logit-transformation for bounded fractional variables (Table S3). For each candidate configuration, we fitted an MLP regression model in *scikit-learn*¹⁰⁹ (Figure 1A4) and optimized the MLP hyperparameters (number and size of hidden layers, activation function, learning-rate strategy, and regularization strength; Table S3) using a nested successive-halving grid search.

For each CV split, we evaluated the resulting model on the calibration set using the root-mean-square error (RMSE; Figure 1A6). This value served as the objective function for the Bayesian optimization framework OPTUNA¹¹⁰, which jointly optimized the parameters listed in Table S3. We used 4,000 random-sampling iterations followed by 4,000 iterations using a tree-structured Parzen estimator¹¹¹ for parameter sampling (Figure 1A7). Each iteration generated a new simulated training dataset and corresponding MLP inversion model. We retained the model with lowest calibration RMSE from each CV split (Figure 1A8) and combined these five selected models into the S2BIOPHYS ensemble (Figure 1A9).

Model Calibration / Validation

We used in-situ reference measurements (RMs) for LAI, FAPAR and FCOVER from the recently published GROUNDED EO⁴¹ RMs dataset, which compiles RMs from three long-term

ecological monitoring networks: NEON based in the US¹¹², ICOS^{113–115} based in Europe, and TERN^{116,117} in Australia (Table S1). At NEON and ICOS sites, LAI, FAPAR and FCOVER were derived from digital hemispherical photography processed with *HemiPy*¹¹⁸, while TERN sites used digital cover photographs processed with *CoverPy*¹¹⁹. Methodological details, including corrections for canopy clumping, missing understory contributions, as well as consistent propagation of measurement uncertainty, are described in Brown et al.⁴¹.

We extracted the temporally closest cloud-free S2 imagery for each RM (Figure 1B2), allowing a maximum temporal offset of +/- 15 days (median temporal difference 2.56 days). We computed mean reflectance within a 20 x 20 m (NEON, ICOS) or 100 x 100 m (TERN) window centered on the RM location, following network-specific protocols^{115,120,121}. After spatio-temporal overlay, a total of 11,038 RMs were retained, representing 1,385 sampling plots across 71 field sites, 44 ecoregions and nine different biomes, in the US, Europe and Australia (Figure S1, Table S1).

We implemented a spatially blocked five-fold cross-validation (CV) scheme based on ecoregions³⁴ (Figure 1B3). This design evaluates model performance under ecological conditions not encountered during calibration, approximating extrapolation to other ecoregions. We assessed S2BIOPHYS performance using two approaches: (1) ensemble-averaged retrievals (Figure 1B5) and (2) stacking retrievals from each CV fold to form a single out-of-sample (SOS) evaluation set to account for imbalanced folds¹²² (Figure 1B4). The SOS evaluation provides a strictly out-of-sample performance estimate under ecoregion-based spatial blocking. In contrast, the ensemble-averaged approach reflects the operational use of S2BIOPHYS but introduces limited information sharing (data leakage) across folds through PROSAIL parameter optimization and ensemble aggregation. Despite this, ensemble averaging improves retrieval stability and accuracy and is therefore used for comparison with SL2P and GROUNDED-EO GPR, which cannot easily be retrained under the same spatial blocking design.

For both evaluation strategies, we computed standard accuracy metrics, including the coefficient of determination (R^2), mean error (ME), mean absolute error (MAE), root-mean-squared error (RMSE), and normalized RMSE (nRMSE). For land-cover-specific evaluation, we additionally computed a global-mean-referenced coefficient of determination (R_{global}^2), following the general formulation of R^2 , but calculating the total sum of squares within each class relative to the overall variable mean across all RMs rather than the class-specific mean. We also

calculated an uncertainty agreement ratio (UAR), defined as the proportion of retrievals meeting predefined user requirements. We adopted the same user-requirement thresholds as previous studies^{41,69,123} to enable direct comparison, with acceptable retrievals defined by a relative error $\leq 20\%$ or an absolute error ≤ 1.0 for LAle, and a relative error $\leq 20\%$ or an absolute error ≤ 0.1 for FAPAR and FCOVER.

Retrieval Uncertainty and Across-Acquisition Variability

For each pixel and S2 acquisition, we first calculated a raw uncertainty proxy as the sample standard deviation across retrievals from the five independently optimized models, $\sigma_{raw} = std(\hat{y}^{(1)}, \dots, \hat{y}^{(5)})$. This ensemble spread captures model disagreement for a fixed reflectance input but does not account for errors due to reference-measurement uncertainty, S2 observation noise, or imperfect spatiotemporal alignment between RMs and satellite observations. We therefore apply a post-hoc recalibration to align the ensemble spread with the magnitude of retrieval errors observed against GROUNDED-EO RMs (Figure S4). Inspired by a previously proposed linear rescaling⁶⁰, we introduced a smooth scaling function $\tau(\hat{y}_i)$ such that the calibrated within-acquisition uncertainty, $\sigma_{within} = \tau(\hat{y})\sigma_{raw}$, minimizes the mean absolute calibration error (MACE),

$$MACE = \frac{1}{N} \sum_{i=1}^N |e_i - \tau(\hat{y}_i) \sigma_{i,raw}|, \text{ where } e_i = |y_i - \hat{y}_i|.$$

Here, N is the total number of RMs used in this study, and y_i and $\hat{y}_i$ are the reference and retrieved values for pair i , respectively. No independent hold-out set was used to evaluate the uncertainty recalibration and therefore, the reported improvement may be optimistic. We modeled the function $\tau(\hat{y})$ as a cubic B-spline with coefficients optimized using constrained nonlinear minimization. To prevent unstable extrapolation at the limits of the retrieval domain, we imposed fixed boundary conditions by setting $\tau = 1$ at slightly extended endpoints (e.g., at -0.1 and 1.1 for variables bounded in $[0, 1]$).

For pixels with multiple acquisitions T , we quantified across-acquisition variability as the sample standard deviation of the per-acquisition mean retrievals, $\sigma_{across} = std(\bar{\hat{y}}^{(1)}, \dots, \bar{\hat{y}}^{(T)})$, where $\bar{\hat{y}}^{(t)}$ denotes the mean retrieval for a given pixel from acquisition t . Finally, we combined the average within-acquisition uncertainty and the across-acquisition variability to obtain a combined pixel-level uncertainty metric σ_{total} . As a practical approximation, we treated both

components as uncorrelated and calculated the total variance as the sum of the two. Any dependence between the components is not represented and may affect the combined uncertainty estimate. We cannot estimate across-acquisition variability for pixels with only one valid acquisition, so we set it to zero, such that $\sigma_{\text{total}} = \sigma_{\text{within}}$. The per-pixel observation count should therefore be considered when interpreting σ_{total} .

Global-scale retrievals

We accessed atmospherically corrected S2 L2A surface reflectance data³⁶ (Figure 1C1) through GEE, organized in 110 x 110 km tiles in the local UTM projection. We used ten spectral bands: B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12. We produced global maps at 20 m and 100 m spatial resolution. When an S2 band is processed at a spatial resolution different from its native resolution (e.g. 20 m bands at 100 m), GEE resamples pixel values by selecting the next higher-resolution image pyramid, which is precomputed using mean aggregation, followed by nearest-neighbor resampling (e.g., from 80 m to 100 m)¹²⁴.

We applied a multi-stage filtering strategy to reduce the number of input images while retaining high-quality and spectrally consistent observations. First, we removed images where the percentage of cloudy pixels (CPP_{image}) exceeded 70%. We linked the remaining images with the *CloudScorePlus*³⁷ image collection to mask clouds and cloud shadows at the pixel level (Figure 1C2-3). We masked pixels with *CloudScorePlus* quality-assessment scores (*cs* band) below 0.65, where 0 indicates cloudy pixels and 1 indicates clear pixels.

To produce annual global maps, we restricted the input S2 imagery to acquisitions representing near-peak greenness conditions. For each S2 image, we computed mean NDVI ($NDVI_{\text{image}}$) across natural vegetation pixels after excluding defective, snow, water, cropland and other non-natural surface pixels based on the S2 scene classification layer⁴⁷ and the ESA WorldCover map¹¹. We then calculated the 95th percentile of these image-level means ($NDVI_{p95}$) across all images for each orbit, tile, and calendar year (Figure 1C4) and retained images for which $NDVI_{p95} - NDVI_{\text{image}} < 0.15$. As this approach selects S2 observations at the image-level, it provides a consistent and scalable way of identifying periods of near-peak greenness across natural vegetation but does not correspond to pixel- or vegetation-type-specific phenological maxima.

Because orbital overlap increases substantially at high latitudes, we restricted the number of orbits per tile to a maximum of two, following Lang et al.¹², and selected the orbits with the smallest fraction of empty pixels. We then determined the target number of images per tile, orbit, and year from the mean cloudy-pixel proportion of the remaining candidates. We retained four images when mean cloudiness was $\leq 20\%$ and eight when it was $\geq 50\%$; between these thresholds, the number was linearly interpolated and rounded to the nearest integer. These operational bounds balanced cloud-free spatial coverage against processing volume. Candidate images were ranked using a quality score that balances image-level cloudiness and vegetation greenness,

$$q = CPP_{image} + 2 * (1 - NDVI_{image}),$$

and we retained the four to eight images with the lowest scores. For example, in 2022, after cloud-, NDVI-, and orbital filtering, we retained 155,595 S2 images as input for S2BIOPHYS processing, corresponding to approximately 3% of all available S2 images for that year.

We incorporated observation angles from the S2 image metadata as predictors by averaging the viewing incidence zenith and azimuth angles across all bands. Subsequently, we calculated the relative azimuth angle between the sun's azimuth and the mean viewing incidence azimuth angle. Finally, we applied a cosine function to the sun's zenith, the mean viewing zenith and the relative azimuth angle.

While we trained the model locally, we re-implemented the ensemble prediction functionality in GEE (Figure 1C5) to leverage cloud-based computing and direct access to the complete S2 L2A archive through the GEE API¹²⁴. We implemented two custom helper classes that replicate the functionality of the *StandardScaler* and *MLPRegressor* classes in scikit-learn¹⁰⁹, specifically adapted for processing images in GEE.

We performed out-of-range checks on input band reflectances by determining the minimum and maximum values for each S2 band from the LUT before model fitting. We then masked input pixels outside this range, allowing a tolerance of 2% of the reflectance range. Ensemble member retrievals were aggregated to obtain the per-image mean retrieval and the raw ensemble spread (σ_{raw}), which we then recalibrated to the empirical within-acquisition uncertainty (σ_{within}).

In the final aggregation step, we combined retrievals across different acquisitions into a single map (Figure 1C6) by computing the pixel-wise mean, the across-acquisition standard deviation (σ_{across}), and the number of valid observations. We obtained total retrieval uncertainty by combining the within-acquisition uncertainty and across-acquisition variability components as described above. We masked permanent water bodies using the ESA WorldCover v200 map for 2021¹¹ and areas classified as the “Rock and Ice” ecoregion (ECO_ID = 0)³⁴.

The resulting global maps (Figure 1C7) are available in GEE at resolutions of 20 m and 100 m from 2019 to 2025. These images are stored in 6° x 8° grid cells defined by the modified MGRS³⁹ using the original UTM projections of the S2 images. To improve accessibility, we provide annual global mosaics at 100 m and 1 km resolution in the EPSG:4326 coordinate reference system, available on Zenodo as cloud-optimized GeoTIFFs⁴⁰. The pixel values are stored as 16-bit integers with scaling factors of 1,000 for LAI and 10,000 for FAPAR and FCOVER.

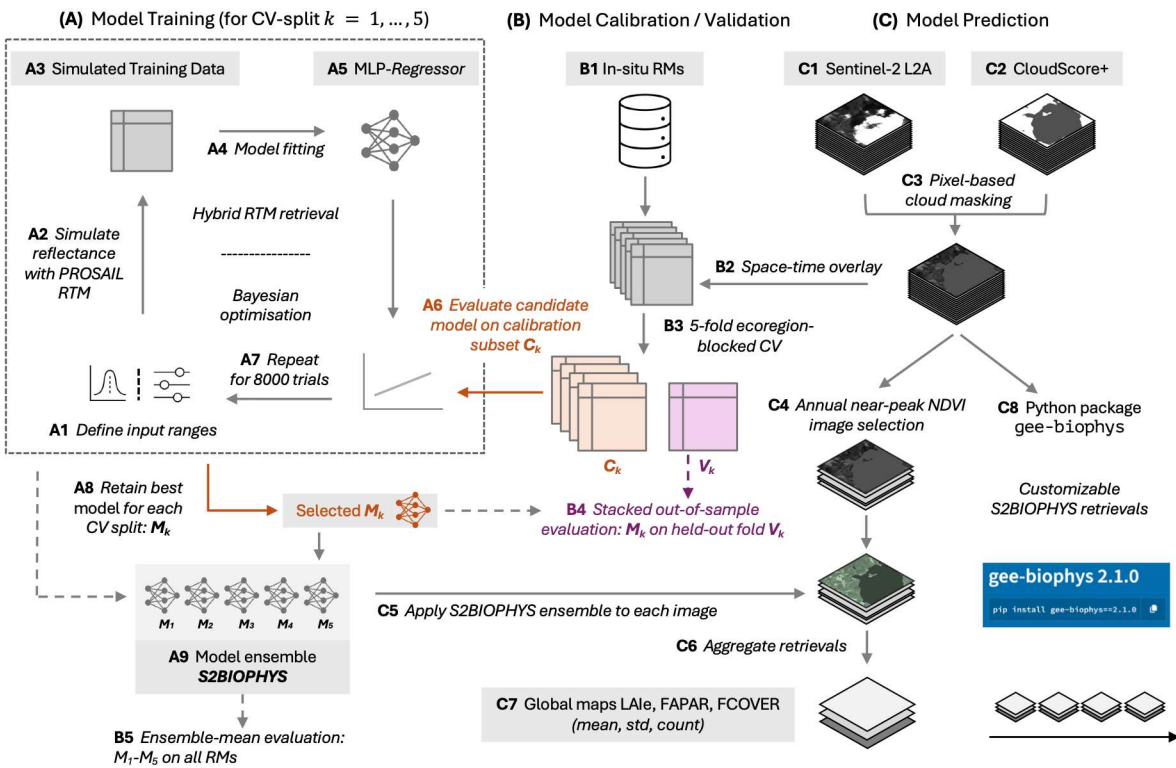
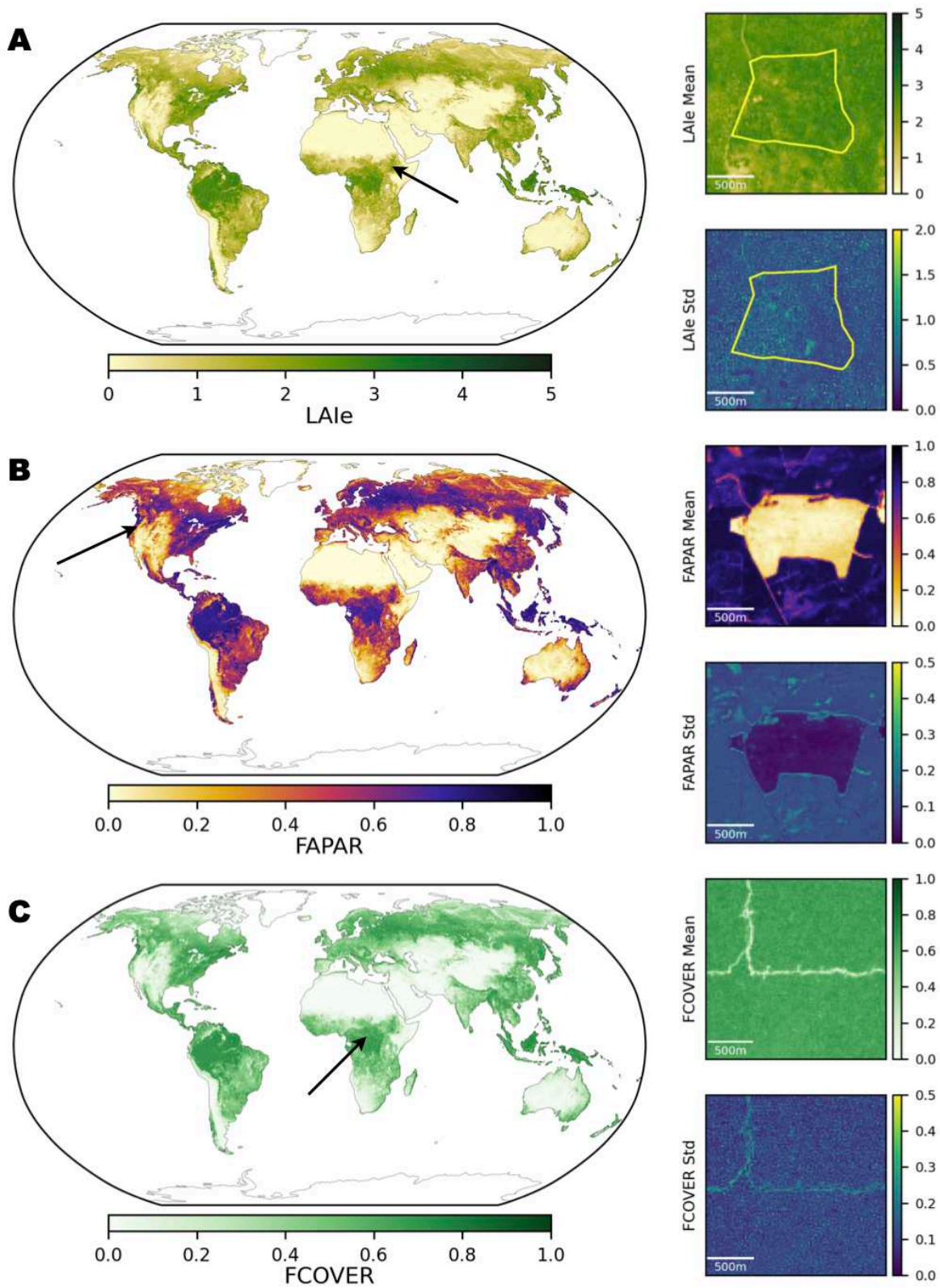
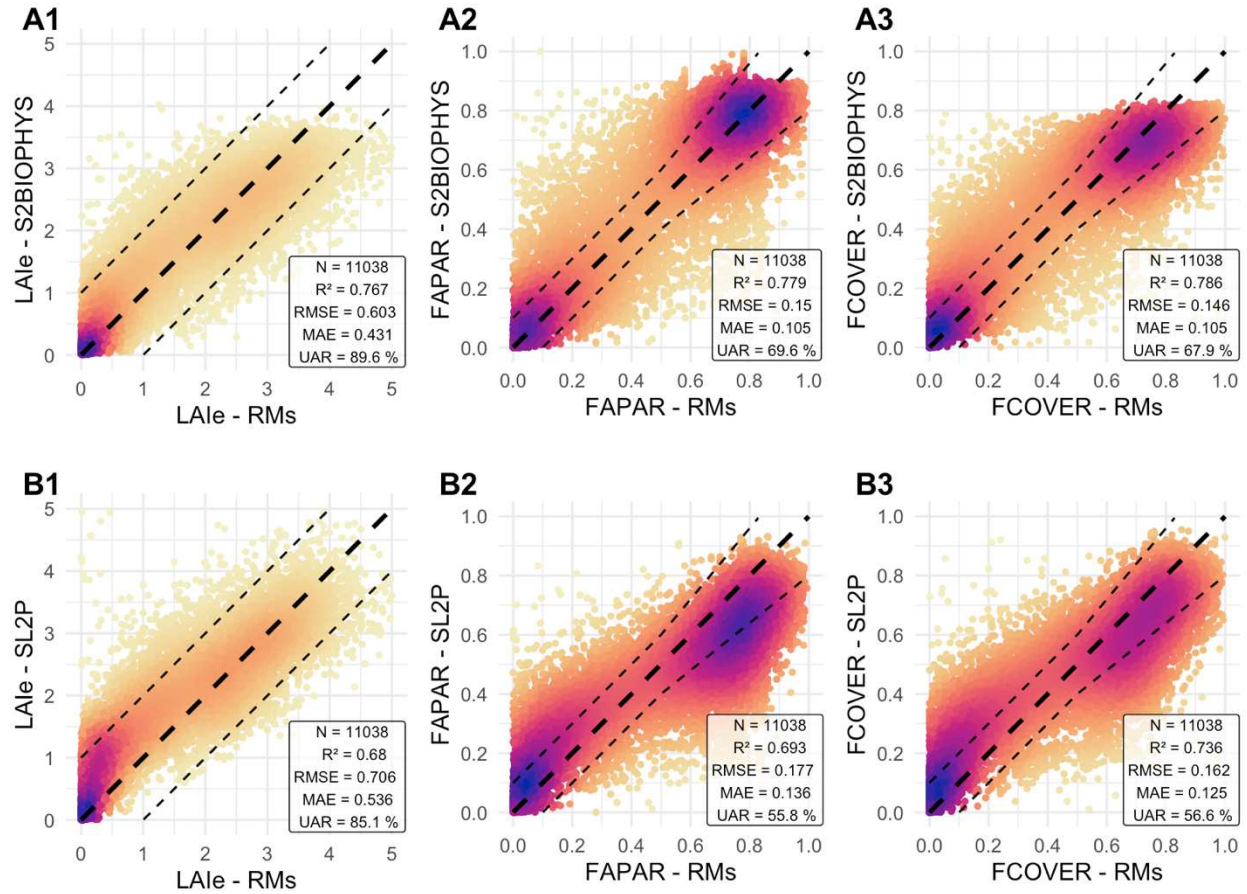


Figure 1 – Model training, calibration, validation, and prediction workflow for S2BIOPHYS. (A) For each of five ecoregion-blocked cross-validation (CV) splits, Bayesian optimization iteratively improved the parametrization (see Table S3) of a hybrid radiative transfer model (RTM) inversion pipeline. We used simulated PROSAIL reflectances to fit each candidate multi-layer perceptron (MLP) model, and in situ reference measurements (RMs) in the corresponding calibration set C_k to evaluate candidate retrieval performance. We retained the best model from each split M_k yielding the five-model S2BIOPHYS ensemble. (B) We paired RMs with cloud-free Sentinel-2 (S2) reflectances. Performance was assessed using stacked out-of-sample retrievals, for which each model M_k was evaluated on its held-out fold V_k , and ensemble-mean retrievals across all five models. (C) S2BIOPHYS produces global annual maps representing near-peak greenness conditions. We applied the full model ensemble to each selected observation and aggregated the resulting retrievals into annual LAI, FAPAR, and FCOVER maps containing the mean, uncertainty estimates, and observation count. Alternatively, the *gee-biophys* Python package supports retrievals for user-defined spatial extents and temporal intervals.



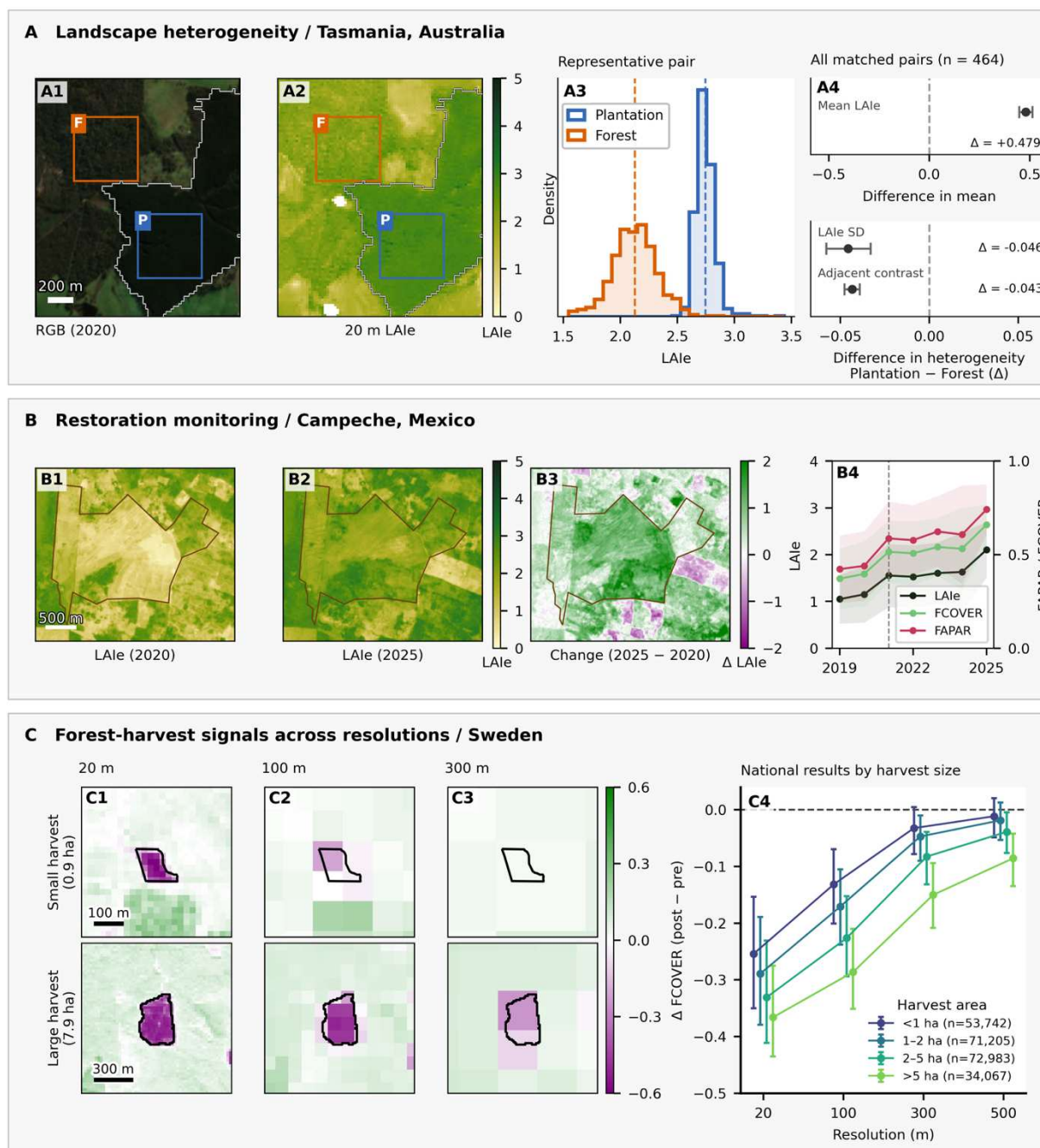
610 **Figure 2 – Global maps of annual LAI, FCOVER and FAPAR.** Retrievals for LAI (**A**),
611 FAPAR (**B**) and FCOVER (**C**) are visualized for the year 2022 in Equal Earth Projection¹²⁵.
612 Local zoom-ins at 20 m resolution show mean and total retrieval uncertainty (std) maps with
613 locations marked by the arrows. (**A**) Agroforestry site in Ethiopia (outlined in yellow) (7.366° N,
614 35.736° E) for 2020. (**B**) Forest clearcut for timber production in Washington State, USA
615 (47.333° N, 123.937° W) for 2024. (**C**) Road network expansion for selective logging in a
616 tropical rainforest in the Congo Basin (1.587° N, 22.149° E) for 2022.

617



618

619 **Figure 3 – Model performance.** Comparison of LAle (A1, B1), FAPAR (A2, B2) and FCOVER
 620 (A3, B3) retrievals between S2BIOPHYS (A1-A3) and SL2P (B1-B3), compared to in-situ
 621 reference measurements (RMs). The thick dashed line represents the 1:1 line and the thinner
 622 dashed lines refer to the uncertainty requirements (see Methods). The annotation box provides
 623 evaluation metrics, including RMSE, MAE, R^2 , UAR and sample size (N). Points in the
 624 scatterplots are colored by the square-root-transformed local density of observations, from
 625 beige (low) to purple (high).



626

627 **Figure 4 – Potential applications of S2BIOPHYS. (A)** Landscape heterogeneity in Tasmania,
 628 Australia. S2 RGB imagery from 2020 **(A1)** and corresponding 20 m LAle 2020 retrievals **(A2)**
 629 are shown for a representative matched plantation-forest pair. Blue and orange boxes identify
 630 the plantation (P) and forest (F) sampling windows, respectively, and the white outline indicates
 631 plantation boundaries¹²⁶. LAle distributions within the two windows are shown in **(A3)**, with
 632 dashed lines indicating the means. **(A4)** shows mean paired differences (plantation minus

forest) across 464 matched pairs, with 95% percentile-bootstrap intervals. See Note S1 for further details. **(B)** Restoration monitoring in Campeche, Mexico. Annual 20 m LAI retrievals are shown for 2020 **(B1)** and 2025 **(B2)**, with the corresponding difference map **(B3)**. Brown outlines indicate the restoration site. **(B4)** shows annual spatial means of LAI, FCOVER, and FAPAR within the restoration site, with shaded bands representing the average pixel-level model uncertainty, and the intervention start year indicated by the dashed line. **(C)** Forest-harvest signals across spatial resolutions in Sweden. FCOVER changes from the year before to the year after harvest for two example harvest areas, displayed at 20 m **(C1)**, 100 m **(C2)**, and 300 m **(C3)**, with harvest boundaries shown in black. **(C4)** shows median FCOVER changes and interquartile ranges by harvest-area class at 20, 100, 300, and 500 m resolutions. See Note S2 for further details.

645 Tables

646 **Table 1 – Model performance.** Comparison of S2BIOPHYS and SL2P retrieval performance
 647 for LAIe, FAPAR and FCOVER. N is the number of reference measurements. For each variable
 648 and NLCD land-cover type⁴⁴, the lower RMSE is highlighted in grey and bold. S2BIOPHYS
 649 performance is reported for the ensemble-averaged retrievals (mean). See Table S4 for
 650 additional evaluation metrics.

Model	NLCD Land cover	N	LAIe		FAPAR		FCOVER	
			RMSE	R ²	RMSE	R ²	RMSE	R ²
S2BIOPHYS	All (mean)	11038	0.60	0.77	0.15	0.78	0.15	0.79
	cultivatedCrops	498	0.50	0.43	0.13	0.78	0.13	0.74
	deciduousForest	3010	0.69	0.72	0.15	0.66	0.15	0.68
	evergreenForest	3286	0.69	0.47	0.19	0.48	0.18	0.51
	grasslandHerbaceous	1685	0.45	0.40	0.12	0.76	0.11	0.78
	mixedForest	501	0.56	0.68	0.10	0.77	0.11	0.78
	pastureHay	519	0.50	0.24	0.14	0.68	0.15	0.66
	shrubScrub	1293	0.31	0.79	0.09	0.79	0.08	0.84
	wetlands	246	0.74	0.58	0.16	0.66	0.16	0.70
SL2P	All	11038	0.70	0.68	0.18	0.69	0.16	0.74
	cultivatedCrops	498	0.80	-0.46	0.13	0.76	0.15	0.64
	deciduousForest	3010	0.66	0.74	0.19	0.48	0.16	0.67
	evergreenForest	3286	0.78	0.32	0.21	0.30	0.20	0.40
	grasslandHerbaceous	1685	0.75	-0.67	0.14	0.68	0.15	0.59
	mixedForest	501	0.52	0.72	0.17	0.31	0.12	0.70
	pastureHay	519	0.82	-0.99	0.14	0.67	0.17	0.58
	shrubScrub	1293	0.43	0.61	0.10	0.74	0.10	0.74
	wetlands	246	0.81	0.49	0.17	0.62	0.17	0.67

651

652 Resource availability

653 *Lead contact*

654 Further information and requests for resources should be directed to and will be fulfilled by the
655 lead contact, Felix Specker (felix.specker@wsl.ch).

656 *Materials availability*

657 No new materials were generated in this study.

658 *Data and code availability*

- 659 - The medium- (100 m) and low-resolution (1 km) global maps are archived on Zenodo
660 and publicly available at <https://doi.org/10.5281/zenodo.19366930>
- 661 - All generated global maps at 20 m and 100 m resolution can be accessed directly on
662 GEE via the following image collections: `ee.ImageCollection("projects/ee-`
663 `speckerfelix/assets/open-earth/{laie, fapar, fcover}_predictions-`
664 `mlp_{100, 20}m_v03")`
- 665 - A public GEE App for exploring the output maps is available at [https://ee-](https://ee-speckerfelix.projects.earthengine.app/view/global-trait-maps)
666 [speckerfelix.projects.earthengine.app/view/global-trait-maps](https://ee-speckerfelix.projects.earthengine.app/view/global-trait-maps).
- 667 - All original code has been deposited at Zenodo at
668 <https://doi.org/10.5281/zenodo.15052995> and is publicly available.
- 669 - An accompanying Python package, *gee-biophys*, for user-defined retrievals is available
670 at <https://pypi.org/project/gee-biophys/>.

671 Acknowledgments

672 The Open-Earth-Monitor Cyberinfrastructure project has received funding from the European
673 Union's Horizon Europe research and innovation programme under grant agreement No.
674 101059548.

675 This study has been undertaken using data from GBOV "Ground Based Observation for
676 Validation" (<https://gbov.land.copernicus.eu/>) founded by European Commission Joint Research
677 Centre under FWC941837, part of the Global Component of the European Union's Copernicus
678 Land Monitoring Service. GBOV products are developed and managed by ACRI-ST, under the

679 coordination of Christophe Lerebourg, with the support from the University of Southampton (J.
680 Dash) and the University of Salford (Luke Brown).

681 We would like to thank the ICOS (<https://www.icos-cp.eu>) PIs (Ivan Janssens, Ivan Mammarella,
682 Caroline Vincke, Nicola Arriga, Jean-Christophe Domec, Michal Heliasz, Anke Hildebrandt,
683 Daniel Berveiller, Alexander Knohl) for providing the digital hemispherical photography dataset
684 over Europe used by GBOV to generate the reference measurements used in this paper.

685 We thank Courtney Meier (NEON) for providing the digital hemispherical photography dataset
686 over North America used by GBOV to generate the reference measurements used in this paper.
687 The National Ecological Observatory Network (NEON) is a program sponsored by the National
688 Science Foundation and operated under cooperative agreement by Battelle. This material is
689 based in part upon work supported by the National Science Foundation through the
690 NEON Program.

691 We thank Ernesto Lopez-Baeza and William Woodgate for supporting the GBOV ground
692 instrumentation source of digital hemispherical photography data used in this paper over the
693 Tumbarumba and Valencia sites.

694 Anna K. Schweiger acknowledges support through NASA-ROSES, Grant No.
695 80NSSC24K0866. Bert Gielen acknowledges the support of the Research Foundation Flanders
696 (FWO) for the ICOS research infrastructure and the EU project NUBICOS (GA No. 101130676).

697 Luke Brown acknowledges that this activity was partly carried out under the Living Planet
698 Fellowship, a programme funded by the European Space Agency (ESA). The view expressed in
699 this publication can in no way be taken to reflect the official opinion of the ESA.

700 We sincerely thank the Google Earth Engine team and Noel Gorelick for their invaluable support
701 in enabling large-scale processing and providing the necessary computational and storage
702 resources.

703 *Author Contributions*

704 Conceptualization, A.K.S., J.H. and T.W.C.; methodology, F.S., A.K.S. and J.H.; data curation,
705 F.S., L.A.B, J.D, R.G, G.B.; analysis, modeling and code development, F.S.; supervision, J.H.
706 and T.W.C.; writing – original draft, F.S. and J.H; writing – review & editing, all authors.

707 *Declaration of interests*

708 The authors declare no competing interests.

709 *Declaration of generative AI and AI-assisted technologies in the writing process*

710 During the preparation of this work the authors used ChatGPT to improve the readability and
711 language of the manuscript. After using this tool, the authors reviewed and edited the text and
712 they take full responsibility for the content of the published article.

713 **Supplemental information**

714 Document S1. Supplemental Tables S1–S6; Figures S1–S9; Notes S1–S4.

715 **Bibliography**

- 716 1. Adla, K., Dejan, K., Neira, D., and Dragana, Š. (2022). Degradation of ecosystems and loss
717 of ecosystem services. In *One Health* (Elsevier), pp. 281–327.
718 <https://doi.org/10.1016/B978-0-12-822794-7.00008-3>.
- 719 2. Myers, S.S., Gaffikin, L., Golden, C.D., Ostfeld, R.S., H. Redford, K., H. Ricketts, T.,
720 Turner, W.R., and Osofsky, S.A. (2013). Human health impacts of ecosystem alteration.
721 *Proc. Natl. Acad. Sci. U.S.A.* *110*, 18753–18760. <https://doi.org/10.1073/pnas.1218656110>.
- 722 3. Sutton, P.C., Anderson, S.J., Costanza, R., and Kubiszewski, I. (2016). The ecological
723 economics of land degradation: Impacts on ecosystem service values. *Ecological*
724 *Economics* *129*, 182–192. <https://doi.org/10.1016/j.ecolecon.2016.06.016>.
- 725 4. Mo, L., Zohner, C.M., Reich, P.B., Liang, J., De Miguel, S., Nabuurs, G.-J., Renner, S.S.,
726 Van Den Hoogen, J., Araza, A., Herold, M., et al. (2023). Integrated global assessment of
727 the natural forest carbon potential. *Nature* *624*, 92–101. [https://doi.org/10.1038/s41586-](https://doi.org/10.1038/s41586-023-06723-z)
728 [023-06723-z](https://doi.org/10.1038/s41586-023-06723-z).
- 729 5. Dodds, W.K., Perkin, J.S., and Gerken, J.E. (2013). Human Impact on Freshwater Ecosystem
730 Services: A Global Perspective. *Environ. Sci. Technol.* *47*, 9061–9068.
731 <https://doi.org/10.1021/es4021052>.

6. Malmqvist, B., and Rundle, S. (2002). Threats to the running water ecosystems of the world. *Envir. Conserv.* 29, 134–153. <https://doi.org/10.1017/S0376892902000097>.
7. Gann, G.D., McDonald, T., Walder, B., Aronson, J., Nelson, C.R., Jonson, J., Hallett, J.G., Eisenberg, C., Guariguata, M.R., Liu, J., et al. (2019). International principles and standards for the practice of ecological restoration. Second edition. *Restoration Ecology* 27. <https://doi.org/10.1111/rec.13035>.
8. UN. General Assembly (2019). United Nations Decade on Ecosystem Restoration (2021–2030), Resolution A/RES/73/284.
9. de Almeida, D.R.A., Stark, S.C., Valbuena, R., Broadbent, E.N., Silva, T.S.F., de Resende, A.F., Ferreira, M.P., Cardil, A., Silva, C.A., Amazonas, N., et al. (2020). A new era in forest restoration monitoring. *Restoration Ecology* 28, 8–11. <https://doi.org/10.1111/rec.13067>.
10. Brown, C.F., Brumby, S.P., Guzder-Williams, B., Birch, T., Hyde, S.B., Mazzariello, J., Czerwinski, W., Pasquarella, V.J., Haertel, R., Ilyushchenko, S., et al. (2022). Dynamic World, Near real-time global 10 m land use land cover mapping. *Sci Data* 9, 251. <https://doi.org/10.1038/s41597-022-01307-4>.
11. Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., et al. (2022). ESA WorldCover 10 m 2021 v200. Version v200 (Zenodo). <https://doi.org/10.5281/zenodo.7254221>
12. Lang, N., Jetz, W., Schindler, K., and Wegner, J.D. (2023). A high-resolution canopy height model of the Earth. *Nat Ecol Evol* 7, 1779–1789. <https://doi.org/10.1038/s41559-023-02206-6>.
13. Weiss, M. (2020). S2ToolBox Level 2 products: LAI, FAPAR, FCOVER (Version 2.1). http://step.esa.int/docs/extra/ATBD_S2ToolBox_V2.1.pdf.
14. Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G.P.W., and Moreno, J. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties – A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 108, 273–290. <https://doi.org/10.1016/j.isprsjprs.2015.05.005>.
15. Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., and Ustin, S.L. (2009). PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment* 113, S56–S66. <https://doi.org/10.1016/j.rse.2008.01.026>.
16. Hauser, L.T., Féret, J.-B., An Binh, N., Van Der Windt, N., Sil, Â.F., Timmermans, J., Soudzilovskaia, N.A., and Van Bodegom, P.M. (2021). Towards scalable estimation of plant functional diversity from Sentinel-2: In-situ validation in a heterogeneous (semi-)natural landscape. *Remote Sensing of Environment* 262, 112505. <https://doi.org/10.1016/j.rse.2021.112505>.

17. Binh, N.A., Hauser, L.T., Viet Hoa, P., Thi Phuong Thao, G., An, N.N., Nhut, H.S., Phuong, T.A., and Verrelst, J. (2022). Quantifying mangrove leaf area index from Sentinel-2 imagery using hybrid models and active learning. *International Journal of Remote Sensing* 43, 5636–5657. <https://doi.org/10.1080/01431161.2021.2024912>.
18. Kganyago, M., Mhangara, P., Alexandridis, T., Laneve, G., Ovakoglou, G., and Mashiyi, N. (2020). Validation of sentinel-2 leaf area index (LAI) product derived from SNAP toolbox and its comparison with global LAI products in an African semi-arid agricultural landscape. *Remote Sensing Letters* 11, 883–892. <https://doi.org/10.1080/2150704X.2020.1767823>.
19. Fernandes, R., Brown, L., Canisius, F., Dash, J., He, L., Hong, G., Huang, L., Le, N.Q., MacDougall, C., Meier, C., et al. (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 Sensing of Environment* 293, 113600. <https://doi.org/10.1016/j.rse.2023.113600>.
20. Fernandes, R., Hong, G., Brown, L.A., Dash, J., Harvey, K., Kalimipalli, S., MacDougall, C., Meier, C., Morris, H., Shah, H., et al. (2024). Not just a pretty picture: Mapping Leaf Area Index at 10 m resolution using Sentinel-2. *Remote Sensing of Environment* 311, 114269. <https://doi.org/10.1016/j.rse.2024.114269>.
21. Wan, L., Ryu, Y., Dechant, B., Hwang, Y., Feng, H., Kang, Y., Jeong, S., Lee, J., Choi, C., and Bae, J. (2024). Correcting confounding canopy structure, biochemistry and soil background effects improves leaf area index estimates across diverse ecosystems from Sentinel-2 imagery. *Remote Sensing of Environment* 309, 114224. <https://doi.org/10.1016/j.rse.2024.114224>.
22. Hu, Q., Yang, J., Xu, B., Huang, J., Memon, M.S., Yin, G., Zeng, Y., Zhao, J., and Liu, K. (2020). Evaluation of Global Decametric-Resolution LAI, FAPAR and FVC Estimates Derived from Sentinel-2 Imagery. *Remote Sensing* 12, 912. <https://doi.org/10.3390/rs12060912>.
23. Brown, L.A., Fernandes, R., Djamai, N., Meier, C., Gobron, N., Morris, H., Canisius, F., Bai, G., Lerebourg, C., Lanconelli, C., et al. (2021). Validation of baseline and modified Sentinel-2 Level 2 Prototype Processor leaf area index retrievals over the United States. *ISPRS Journal of Photogrammetry and Remote Sensing* 175, 71–87. <https://doi.org/10.1016/j.isprsjprs.2021.02.020>.
24. Skidmore, A.K., Coops, N.C., Neinavaz, E., Ali, A., Schaepman, M.E., Paganini, M., Kissling, W.D., Vihervaara, P., Darvishzadeh, R., Feilhauer, H., et al. (2021). Priority list of biodiversity metrics to observe from space. *Nat Ecol Evol* 5, 896–906. <https://doi.org/10.1038/s41559-021-01451-x>.
25. Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholczyk, P., and Smets, B. (2013). GEOV1: *LAI* and *FAPAR* essential climate variables and *FCOVER* global time series capitalizing over existing products. Part I: Principles of development and production. *Remote Sensing of Environment* 137, 299–309. <https://doi.org/10.1016/j.rse.2012.12.027>.

26. Zemp, M. (2021). The status of the global climate observing system 2021: The GCOS status report. <https://doi.org/10.5167/UZH-213734>.
27. Fang, H., Baret, F., Plummer, S., and 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>.
28. Chen, J.M., Menges, C.H., and Leblanc, S.G. (2005). Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sensing of Environment* 97, 447–457. <https://doi.org/10.1016/j.rse.2005.05.003>.
29. Chu, D. (2020). Fractional Vegetation Cover. In *Remote Sensing of Land Use and Land Cover in Mountain Region: A Comprehensive Study at the Central Tibetan Plateau*, D. Chu, ed. (Springer), pp. 195–207. https://doi.org/10.1007/978-981-13-7580-4_10.
30. Cai, Y., Zhang, F., Duan, P., Yung Jim, C., Weng Chan, N., Shi, J., Liu, C., Wang, J., Bahtebay, J., and Ma, X. (2022). Vegetation cover changes in China induced by ecological restoration-protection projects and land-use changes from 2000 to 2020. *CATENA* 217, 106530. <https://doi.org/10.1016/j.catena.2022.106530>.
31. Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P. (2012). Consistent assimilation of MERIS FAPAR and atmospheric CO₂ into a terrestrial vegetation model and interactive mission benefit analysis. *Biogeosciences* 9, 3173–3184. <https://doi.org/10.5194/bg-9-3173-2012>.
32. Safanelli, J.L., Hengl, T., Parente, L.L., Minarik, R., Bloom, D.E., Todd-Brown, K., Gholizadeh, A., Mendes, W. de S., and Sanderman, J. (2025). Open Soil Spectral Library (OSSL): Building reproducible soil calibration models through open development and community engagement. *PLOS ONE* 20, e0296545. <https://doi.org/10.1371/journal.pone.0296545>.
33. Green, R.O., Thompson, D.R., and Team, E. (2021). NASA’s Earth Surface Mineral Dust Source Investigation: An Earth Venture Imaging Spectrometer Science Mission. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, pp. 119–122. <https://doi.org/10.1109/IGARSS47720.2021.9554217>.
34. Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N.D., Wikramanayake, E., Hahn, N., Palminteri, S., Hedao, P., Noss, R., et al. (2017). An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *BioScience* 67, 534–545. <https://doi.org/10.1093/biosci/bix014>.
35. Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D’amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., et al. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* 51, 933–938. [https://doi.org/10.1641/0006-3568\(2001\)051%5B0933:TEOTWA%5D2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051%5B0933:TEOTWA%5D2.0.CO;2).

- 846 38. European Space Agency (2021). Copernicus Sentinel-2 MSI Level-2A BOA reflectance
847 product, Collection 1 (processed by ESA). https://doi.org/10.5270/S2_-znk9xsj.
- 848 37. Pasquarella, V.J., Brown, C.F., Czerwinski, W., and Rucklidge, W.J. (2023). Comprehensive
849 Quality Assessment of Optical Satellite Imagery Using Weakly Supervised Video Learning.
850 Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition
851 (CVPR) Workshops, 2125–2135.
- 852 38. Rouse, J.W., Jr., Haas, R.H., Deering, D.W., Schell, J.A., and Harlan, J.C. (1974).
853 Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural
854 Vegetation. Final report. (NASA/GSFC).
- 855 39. European Space Agency (2024). Sentinel-2 Products Specification Document (Issue 15.0,
856 S2-PDGS-CS-DI-PSD).
- 857 40. Specker, F., Schweiger, A.K., Féret, J.-B., Brown, L.A., Lauber, T., Khoury, C.E., Dash, J.,
858 Grousset, R., Gielen, B., Crowther, T.W., et al. (2026). Advancing ecosystem monitoring:
859 Global annual maps of biophysical vegetation properties (LAIe, FAPAR, FCOVER) for
860 2019-2025. (Zenodo). <https://doi.org/10.5281/zenodo.19366930>.
- 861 41. Brown, L.A., Fernandes, R., Verrelst, J., Morris, H., Djamai, N., Reyes-Muñoz, P.,
862 D.Kovács, D., and Meier, C. (2025). GROUNDED EO: Data-driven Sentinel-2 LAI and
863 FAPAR retrieval using Gaussian processes trained with extensive fiducial reference
864 measurements. *Remote Sensing of Environment* 326, 114797.
865 <https://doi.org/10.1016/j.rse.2025.114797>.
- 866 42. Copernicus Land Monitoring Service (2026). Leaf Area Index 2014-present (raster 300 m),
867 global, 10-daily. Version 2. [https://doi.org/10.2909/A696FAB4-C3DC-4DCC-B1D8-](https://doi.org/10.2909/A696FAB4-C3DC-4DCC-B1D8-6FCD9A2CE1EB)
868 [6FCD9A2CE1EB](https://doi.org/10.2909/A696FAB4-C3DC-4DCC-B1D8-6FCD9A2CE1EB).
- 869 43. Copernicus Land Monitoring Service (2026). Fraction of Absorbed Photosynthetically
870 Active Radiation 2014-present (raster 300 m), global, 10-daily. Version 2.
871 <https://doi.org/10.2909/C828C1B2-6EA8-40F3-897D-C92080BA1DAF>.
- 872 44. United States Geological Survey (2026). Annual National Land Cover Database (NLCD)
873 Collection 1 Products. (U.S. Geological Survey). <https://doi.org/10.5066/P94UXNTS>.
- 874 45. GBOV (2023). Algorithm Theoretical Basis Document - Vegetation Products RM4
875 (Tcanopy), RM6 (FIPAR) and RM7 (LAI) and FCOVER (Version 2.2, GBOV-ATBD-
876 RM4-RM6-RM7)
- 877 46. Gorroño, J., Guanter, L., Valentin Graf, L., and Gascon, F. (2024). A Framework for the
878 Estimation of Uncertainties and Spectral Error Correlation in Sentinel-2 Level-2A Data
879 Products. *IEEE Transactions on Geoscience and Remote Sensing* 62, 1–13.
880 <https://doi.org/10.1109/TGRS.2024.3435021>.

- 881 47. European Space Agency (2021). Sentinel-2 Level-2A Algorithm Theoretical Basis
882 Document (ATBD) (Issue 2.10, S2-PDGS-MPC-ATBD-L2A)
- 883 48. Chraïbi, E., de Boissieu, F., Barbier, N., Luque, S., and Féret, J.-B. (2022). Stability in time
884 and consistency between atmospheric corrections: Assessing the reliability of Sentinel-2
885 products for biodiversity monitoring in tropical forests. *International Journal of Applied*
886 *Earth Observation and Geoinformation* 112, 102884.
887 <https://doi.org/10.1016/j.jag.2022.102884>.
- 888 49. Roy, D.P., Li, J., Zhang, H.K., Yan, L., Huang, H., and Li, Z. (2017). Examination of
889 Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a
890 general method to normalize MSI reflectance to nadir BRDF adjusted reflectance. *Remote*
891 *Sensing of Environment* 199, 25–38. <https://doi.org/10.1016/j.rse.2017.06.019>.
- 892 50. Poorter, L., Craven, D., Jakovac, C.C., van der Sande, M.T., Amissah, L., Bongers, F.,
893 Chazdon, R.L., Farrior, C.E., Kambach, S., Meave, J.A., et al. (2021). Multidimensional
894 tropical forest recovery. *Science* 374, 1370–1376. <https://doi.org/10.1126/science.abh3629>.
- 895 51. Heidrich, L., Bae, S., Levick, S., Seibold, S., Weisser, W., Krzystek, P., Magdon, P., Naus, S.,
896 T., Schall, P., Serebryanyk, A., et al. (2020). Heterogeneity–diversity relationships differ
897 between and within trophic levels in temperate forests. *Nat Ecol Evol* 4, 1204–1212.
898 <https://doi.org/10.1038/s41559-020-1245-z>.
- 899 52. Thomsen, M.S., Altieri, A.H., Angelini, C., Bishop, M.J., Bulleri, F., Farhan, R., Frühling, Y.,
900 V.M.M., Gribben, P.E., Harrison, S.B., He, Q., et al. (2022). Heterogeneity within and
901 among co-occurring foundation species increases biodiversity. *Nat Commun* 13, 581.
902 <https://doi.org/10.1038/s41467-022-28194-y>.
- 903 53. Stein, A., Gerstner, K., and Kreft, H. (2014). Environmental heterogeneity as a universal
904 driver of species richness across taxa, biomes and spatial scales. *Ecology Letters* 17, 866–
905 880. <https://doi.org/10.1111/ele.12277>.
- 906 54. Tuanmu, M.-N., and Jetz, W. (2015). A global, remote sensing-based characterization of
907 terrestrial habitat heterogeneity for biodiversity and ecosystem modelling. *Global Ecology*
908 *and Biogeography* 24, 1329–1339. <https://doi.org/10.1111/geb.12365>.
- 909 55. Torresani, M., Rossi, C., Perrone, M., Hauser, L.T., Féret, J.-B., Moudrý, V., Simova, P.,
910 Ricotta, C., Foody, G.M., Kacic, P., et al. (2024). Reviewing the Spectral Variation
911 Hypothesis: Twenty years in the tumultuous sea of biodiversity estimation by remote
912 sensing. *Ecological Informatics* 82, 102702. <https://doi.org/10.1016/j.ecoinf.2024.102702>.
- 913 56. Helfenstein, I.S., Schneider, F.D., Schaepman, M.E., and Morsdorf, F. (2022). Assessing
914 biodiversity from space: Impact of spatial and spectral resolution on trait-based functional
915 diversity. *Remote Sensing of Environment* 275, 113024.
916 <https://doi.org/10.1016/j.rse.2022.113024>.

57. Fernandes, R., Djamai, N., Harvey, K., Hong, G., MacDougall, C., Shah, H., and Sun, L. (2024). Evidence of a bias-variance trade off when correcting for bias in Sentinel 2 forest LAI retrievals using radiative transfer models. *Remote Sensing of Environment* 305, 114060. <https://doi.org/10.1016/j.rse.2024.114060>.
58. European Union's Copernicus Land Monitoring Service information Leaf Area Index 2016-present (raster 10 m), Europe, daily. <https://land.copernicus.eu/en/products/vegetation/high-resolution-leaf-area-index>.
59. Kang, Y., Ozdogan, M., Gao, F., Anderson, M.C., White, W.A., Yang, Y., Yang, Y., and Erickson, T.A. (2021). A data-driven approach to estimate leaf area index for Landsat images over the contiguous US. *Remote Sensing of Environment* 258, 112383. <https://doi.org/10.1016/j.rse.2021.112383>.
60. Martínez-Ferrer, L., Moreno-Martínez, Á., Campos-Taberner, M., García-Haro, F.J., Muñoz-Mari, J., Running, S.W., Kimball, J., Clinton, N., and Camps-Valls, G. (2022). Quantifying uncertainty in high resolution biophysical variable retrieval with machine learning. *Remote Sensing of Environment* 280, 113199. <https://doi.org/10.1016/j.rse.2022.113199>.
61. Myneni, R., Knyazikhin, Y., and Park, T. (2021). MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V061. (NASA EOSDIS Land Processes Distributed Active Archive Center). <https://doi.org/10.5067/MODIS/MOD15A2H.061>.
62. Copernicus Land Monitoring Service (2026). Fraction of Vegetation Cover 2014-present (raster 300 m), global, 10-daily. Version 2. <https://doi.org/10.2909/E7B99351-7A89-42C7-8BBA-1D9FC0463946>.
63. Samberg, L.H., Gerber, J.S., Ramankutty, N., Herrero, M., and West, P.C. (2016). Subnational distribution of average farm size and smallholder contributions to global food production. *Environ. Res. Lett.* 11, 124010. <https://doi.org/10.1088/1748-9326/11/12/124010>.
64. Fan, S., and Rue, C. (2020). The Role of Smallholder Farms in a Changing World. In *The Role of Smallholder Farms in Food and Nutrition Security*, S. Gomez Y Paloma, L. Riesgo, and K. Louhichi, eds. (Springer International Publishing), pp. 13–28. https://doi.org/10.1007/978-3-030-42148-9_2.
65. Ricciardi, V., Mehrabi, Z., Wittman, H., James, D., and Ramankutty, N. (2021). Higher yields and more biodiversity on smaller farms. *Nat Sustain* 4, 651–657. <https://doi.org/10.1038/s41893-021-00699-2>.
66. Shyamsundar, P., Cohen, F., Boucher, T.M., Kroeger, T., Erbaugh, J.T., Waterfield, G., Clarke, C., Cook-Patton, S.C., Garcia, E., Juma, K., et al. (2022). Scaling smallholder tree cover restoration across the tropics. *Global Environmental Change* 76, 102591. <https://doi.org/10.1016/j.gloenvcha.2022.102591>.

67. Kovács, D.D., De Clerck, E., and Verrelst, J. (2025). PyEOGPR: A Python package for vegetation trait mapping with Gaussian Process Regression on Earth observation cloud platforms. *Ecological Informatics* 92, 103497. <https://doi.org/10.1016/j.ecoinf.2025.103497>.
68. Djamai, N., Fernandes, R., Sun, L., Canisius, F., and Hong, G. (2024). Python version of Simplified Level 2 Prototype Processor for Retrieving Canopy Biophysical Variables from Sentinel 2 Multispectral Instrument Data. <https://doi.org/10.5281/ZENODO.10654520>.
69. Djamai, N., Fernandes, R., Weiss, M., McNairn, H., and Goïta, 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 Sensing of Environment* 225, 416–430. <https://doi.org/10.1016/j.rse.2019.03.020>.
70. Verrelst, J., Malenovský, Z., Van der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J.-P., Lewis, P., North, P., and Moreno, J. (2019). Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods. *Surv Geophys* 40, 589–629. <https://doi.org/10.1007/s10712-018-9478-y>.
71. Wang, J., Lopez-Lozano, R., Weiss, M., Buis, S., Li, W., Liu, S., Baret, F., and Zhang, J. (2022). Crop specific inversion of PROSAIL to retrieve green area index (GAI) from several decametric satellites using a Bayesian framework. *Remote Sensing of Environment* 278, 113085. <https://doi.org/10.1016/j.rse.2022.113085>.
73. Friedl, M., and Sulla-Menashe, D. (2022). MODIS/terra+aqua land cover type yearly L3 global 500m SIN grid V061. (NASA Land Processes Distributed Active Archive Center). <https://doi.org/10.5067/MODIS/MCD12Q1.061>.
73. Buchhorn, M., Smets, B., Bertels, L., Roo, B.D., Lesiv, M., Tsendbazar, N.-E., Li, L., and Tarko, A. (2021). Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual (Zenodo) <https://doi.org/10.5281/zenodo.4723921>.
74. Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J.C., Mathis, M., and Brumby, S.P. (2021). Global land use / land cover with Sentinel 2 and deep learning. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, pp. 4704–4707. <https://doi.org/10.1109/IGARSS47720.2021.9553499>.
75. Brown, C.F., Brumby, S.P., Guzder-Williams, B., Birch, T., Hyde, S.B., Mazzariello, J., Czerwinski, W., Pasquarella, V.J., Haertel, R., Ilyushchenko, S., et al. (2022). Dynamic World, Near real-time global 10 m land use land cover mapping. *Sci Data* 9, 251. <https://doi.org/10.1038/s41597-022-01307-4>.
76. Fang, H., Li, W., and Myneni, R.B. (2013). The Impact of Potential Land Cover Misclassification on MODIS Leaf Area Index (LAI) Estimation: A Statistical Perspective. *Remote Sensing* 5, 830–844. <https://doi.org/10.3390/rs5020830>.

989 77. Yang, L., Jia, K., Liang, S., Wei, X., Yao, Y., and Zhang, X. (2017). A Robust Algorithm for
990 Estimating Surface Fractional Vegetation Cover from Landsat Data. *Remote Sensing* 9,
991 857. <https://doi.org/10.3390/rs9080857>.

992 78. Braghieri, R.K., Wang, Y., Gagné-Landmann, A., Brodrick, P.G., Bloom, A.A., Norton,
993 A.J., Ma, S., Levine, P., Longo, M., Deck, K., et al. (2023). The Importance of
994 Hyperspectral Soil Albedo Information for Improving Earth System Model Projections.
995 *AGU Advances* 4, e2023AV000910. <https://doi.org/10.1029/2023AV000910>.

996 79. De Clerck, E., D.Kovács, D., Berger, K., Schlerf, M., and Verrelst, J. (2024). Optimizing
997 hybrid models for canopy nitrogen mapping from Sentinel-2 in Google Earth Engine.
998 *ISPRS Journal of Photogrammetry and Remote Sensing* 218, 530–545.
999 <https://doi.org/10.1016/j.isprsjprs.2024.11.005>.

1000 80. Estévez, J., Salinero-Delgado, M., Berger, K., Pipia, L., Rivera-Caicedo, J.P., Wocher, M.,
1001 Reyes-Muñoz, P., Tagliabue, G., Boschetti, M., and Verrelst, J. (2022). Gaussian processes
1002 retrieval of crop traits in Google Earth Engine based on Sentinel-2 top-of-atmosphere data.
1003 *Remote Sensing of Environment* 273, 112958. <https://doi.org/10.1016/j.rse.2022.112958>.

1004 81. Kovács, D.D., Reyes-Muñoz, P., Salinero-Delgado, M., Mészáros, V.I., Berger, K., and
1005 Verrelst, J. (2023). Cloud-Free Global Maps of Essential Vegetation Traits Processed from
1006 the TOA Sentinel-3 Catalogue in Google Earth Engine. *Remote Sensing* 15, 3404.
1007 <https://doi.org/10.3390/rs15133404>.

1008 82. Baret, F., Jacquemoud, S., Guyot, G., and Leprieux, C. (1992). Modeled analysis of the
1009 biophysical nature of spectral shifts and comparison with information content of broad
1010 bands. *Remote Sensing of Environment* 41, 133–142. [https://doi.org/10.1016/0034-](https://doi.org/10.1016/0034-4257(92)90073-S)
1011 [4257\(92\)90073-S](https://doi.org/10.1016/0034-4257(92)90073-S).

1012 83. Campos-Taberner, M., Moreno-Martínez, Á., García-Haro, F., Camps-Valls, G., Robinson,
1013 N., Kattge, J., and Running, S. (2018). Global Estimation of Biophysical Variables from
1014 Google Earth Engine Platform. *Remote Sensing* 10, 1167.
1015 <https://doi.org/10.3390/rs10081167>.

1016 84. Zérah, Y., Valero, S., and Inglada, J. (2024). Physics-constrained deep learning for
1017 biophysical parameter retrieval from Sentinel-2 images: Inversion of the PROSAIL model.
1018 *Remote Sensing of Environment* 312, 114309. <https://doi.org/10.1016/j.rse.2024.114309>.

1019 85. National Ecological Observatory Network (NEON) (2023). Plant foliar traits
1020 (DP1.10026.001): RELEASE-2023. 81.4 MB. <https://doi.org/10.48443/tmrs-fb32>.

1021 86. Masarotto, G., and Varin, C. (2017). Gaussian Copula Regression in R. *J. Stat. Soft.* 77.
1022 <https://doi.org/10.18637/jss.v077.i08>.

1023 87. Féret, J.-B., and De Boissieu, F. (2024). prosail: PROSAIL leaf and canopy radiative transfer
1024 model and inversion routines. Version 2.4.1. <https://github.com/jbferet/prosail>

- 1025 88. Thompson, D.R., Brodrick, P.G., Green, R.O., Kalashnikova, O., Okin, G., Olson-Duvall,
1026 W., and Painter, T. (2020). EMIT L2A Algorithm: Surface Reflectance and Scene Content
1027 Masks - Theoretical Basis.
- 1028 89. Green, R. (2022). EMIT L2A Estimated Surface Reflectance and Uncertainty and Masks 60
1029 m V001. <https://doi.org/https://doi.org/10.5067/EMIT/EMITL2ARFL.001>.
- 1030 90. Safanelli, J.L., Hengl, T., Sanderman, J., and Parente, L. (2021). Open Soil Spectral Library
1031 (training data and calibration models). Version v1.2 (Zenodo).
1032 <https://doi.org/10.5281/zenodo.7599269>.
- 1033 91. Agroforestry (ICRAF), W., and Centre (ISRIC), I.S.R. and I. (2021). ICRAF-ISRIC Soil
1034 VNIR Spectral Library. (World Agroforestry (ICRAF)).
1035 <https://doi.org/10.34725/DVN/MFHA9C>.
- 1036 92. Aitkenhead, M.J., and Black, H.I.J. (2018). Exploring the Impact of Different Input Data
1037 Types on Soil Variable Estimation Using the ICRAF-ISRIC Global Soil Spectral Database.
1038 *Appl Spectrosc* 72, 188–198. <https://doi.org/10.1177/0003702817739013>.
- 1039 93. Garrett, L.G., Sanderman, J., Palmer, D.J., Dean, F., Patel, S., Bridson, J.H., and Carlin, T.
1040 (2022). Mid-infrared spectroscopy for planted forest soil and foliage nutrition predictions,
1041 New Zealand case study. *Trees, Forests and People* 8, 100280.
1042 <https://doi.org/10.1016/j.tfp.2022.100280>.
- 1043 94. Jović, B., Ćirić, V., Kovačević, M., Šeremešić, S., and Kordić, B. (2019). Empirical equation
1044 for preliminary assessment of soil texture. *Spectrochimica Acta Part A: Molecular and*
1045 *Biomolecular Spectroscopy* 206, 506–511. <https://doi.org/10.1016/j.saa.2018.08.039>.
- 1046 95. Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., and Fernández-Ugalde, O. (2018). LUCAS
1047 Soil, the largest expandable soil dataset for Europe: a review. *European Journal of Soil*
1048 *Science* 69, 140–153. <https://doi.org/10.1111/ejss.12499>.
- 1049 96. Sanderman, J., Savage, K., and Dangal, S.R.S. (2020). Mid-infrared spectroscopy for
1050 prediction of soil health indicators in the United States. *Soil Science Society of America*
1051 *Journal* 84, 251–261. <https://doi.org/10.1002/saj2.20009>.
- 1052 97. Schiedung, M., Bellè, S.-L., Malhotra, A., and Abiven, S. (2022). Organic carbon stocks,
1053 quality and prediction in permafrost-affected forest soils in North Canada. *CATENA* 213,
1054 106194. <https://doi.org/10.1016/j.catena.2022.106194>.
- 1055 98. Summerauer, L., Baumann, P., Ramirez-Lopez, L., Barthel, M., Bauters, M., Bukombe, B.,
1056 Reichenbach, M., Boeckx, P., Kearsley, E., Van Oost, K., et al. (2021). The central African
1057 soil spectral library: a new soil infrared repository and a geographical prediction analysis.
1058 *SOIL* 7, 693–715. <https://doi.org/10.5194/soil-7-693-2021>.
- 1059 99. Vågen, T.-G., Winowiecki, L.A., Desta, L., Tondoh, E.J., Weullow, E., Shepherd, K., and
1060 Sila, A. (2020). Mid-Infrared Spectra (MIRS) from ICRAF Soil and Plant Spectroscopy

1061 Laboratory: Africa Soil Information Service (AfSIS) Phase I 2009-2013. (World
1062 Agroforestry (ICRAF)). <https://doi.org/10.34725/DVN/QXCWP1>.

1063 100. Wijewardane, N.K., Ge, Y., Wills, S., and Libohova, Z. (2018). Predicting Physical and
1064 Chemical Properties of US Soils with a Mid-Infrared Reflectance Spectral Library. *Soil*
1065 *Science Society of America Journal* 82, 722–731.
1066 <https://doi.org/10.2136/sssaj2017.10.0361>.

1067 101. Wijewardane, N.K., Ge, Y., Wills, S., and Loecke, T. (2016). Prediction of Soil Carbon in
1068 the Conterminous United States: Visible and Near Infrared Reflectance Spectroscopy
1069 Analysis of the Rapid Carbon Assessment Project. *Soil Science Society of America Journal*
1070 80, 973–982. <https://doi.org/10.2136/sssaj2016.02.0052>.

1071 102. Wills, S., Loecke, T., Sequeira, C., Teachman, G., Grunwald, S., and West, L.T. (2014).
1072 Overview of the U.S. Rapid Carbon Assessment Project: Sampling Design, Initial Summary
1073 and Uncertainty Estimates. In *Soil Carbon*, A. E. Hartemink and K. McSweeney, eds.
1074 (Springer International Publishing), pp. 95–104. [https://doi.org/10.1007/978-3-319-04084-](https://doi.org/10.1007/978-3-319-04084-4_10)
1075 [4_10](https://doi.org/10.1007/978-3-319-04084-4_10).

1076 103. European Space Agency (2024). Sentinel-2 spectral response functions (S2-SRF)
1077 (European Space Agency). Version 4.0, COPE-GSEG-EOPG-TN-15-0007

1078 104. Jacquemoud, S., and Baret, F. (1990). PROSPECT: A model of leaf optical properties
1079 spectra. *Remote Sensing of Environment* 34, 75–91. [https://doi.org/10.1016/0034-](https://doi.org/10.1016/0034-4257(90)90100-Z)
1080 [4257\(90\)90100-Z](https://doi.org/10.1016/0034-4257(90)90100-Z).

1081 105. Verhoef, W. (1984). Light scattering by leaf layers with application to canopy reflectance
1082 modeling: The SAIL model. *Remote Sensing of Environment* 16, 125–141.
1083 [https://doi.org/10.1016/0034-4257\(84\)90057-9](https://doi.org/10.1016/0034-4257(84)90057-9).

1084 106. Féret, J.-B., Gitelson, A.A., Noble, S.D., and Jacquemoud, S. (2017). PROSPECT-D:
1085 Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of*
1086 *Environment* 193, 204–215. <https://doi.org/10.1016/j.rse.2017.03.004>.

1087 107. Féret, J.-B., and De Boissieu, F. (2024). prospect: an R package to link leaf optical
1088 properties with their chemical and structural properties with the leaf model PROSPECT.
1089 *JOSS* 9, 6027. <https://doi.org/10.21105/joss.06027>.

1090 108. Köhler, P., Guanter, L., Kobayashi, H., Walther, S., and Yang, W. (2018). Assessing the
1091 potential of sun-induced fluorescence and the canopy scattering coefficient to track large-
1092 scale vegetation dynamics in Amazon forests. *Remote Sensing of Environment* 204, 769–
1093 785. <https://doi.org/10.1016/j.rse.2017.09.025>.

1094 109. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel,
1095 M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine Learning
1096 in Python. *Journal of Machine Learning Research* 12, 2825–2830.

- 1097 110. Akiba, T., Sano, S., Yanase, T., Ohta, T., and Masanori Koyama (2019). Optuna: A Next-
 1098 generation Hyperparameter Optimization Framework. In Proceedings of the 25th ACM
 1099 SIGKDD International Conference on Knowledge Discovery & Data Mining KDD '19.
 1100 (Association for Computing Machinery), pp. 2623–2631.
 1101 <https://doi.org/10.1145/3292500.3330701>.
- 1102 111. Bergstra, J., Bardenet, R., Bengio, Y., and Kégl, B. (2011). Algorithms for hyper-parameter
 1103 optimization. In Advances in neural information processing systems (Curran Associates,
 1104 Inc.).
- 1105 112. National Ecological Observatory Network (NEON) (2024). Digital hemispheric photos of
 1106 plot vegetation (DP1.10017.001): RELEASE-2024.
 1107 <https://doi.org/https://doi.org/10.48443/jbb9-2t38>.
- 1108 113. Heiskanen, J., Brümmer, C., Buchmann, N., Calfapietra, C., Chen, H., Gielen, B.,
 1109 Gkritzalis, T., Hammer, S., Hartman, S., Herbst, M., et al. (2022). The Integrated Carbon
 1110 Observation System in Europe. *Bulletin of the American Meteorological Society* 103,
 1111 E855–E872. <https://doi.org/10.1175/BAMS-D-19-0364.1>.
- 1112 114. ICOS RI, Aalto, J., Aalto, P., Aaltonen, H., Aiguier, T., Akubia, J., Ala-Könni, J.,
 1113 Alivernini, A., Aluome, C., Andersson, T., et al. (2024). Ecosystem final quality (L2)
 1114 product in ETC-Archive format - release 2024-1. Preprint at ICOS ERIC - Carbon Portal,
 1115 <https://doi.org/10.18160/G5KZ-ZD83>.
- 1116 115. Gielen, B., Acosta, M., Altimir, N., Buchmann, N., Cescatte, A., Ceschia, E., Fleck, S.,
 1117 Hortnagal, L., Klumpp, K., Kolari, P., et al. (2018). Ancillary vegetation measurements at
 1118 ICOS ecosystem stations. <https://doi.org/10.1515/intag-2017-0048>.
- 1119 116. Karan, M., Liddell, M., Prober, S.M., Arndt, S., Beringer, J., Boer, M., Cleverly, J., Eamus,
 1120 D., Grace, P., Van Gorsel, E., et al. (2016). The Australian SuperSite Network: A
 1121 continental, long-term terrestrial ecosystem observatory. *Science of The Total Environment*
 1122 568, 1263–1274. <https://doi.org/10.1016/j.scitotenv.2016.05.170>.
- 1123 117. Cleverly, J., Eamus, D., Edwards, W., Grant, M., Grundy, M.J., Held, A., Karan, M., Lowe,
 1124 A.J., Prober, S.M., Sparrow, B., et al. (2019). TERN, Australia's land observatory:
 1125 Addressing the global challenge of forecasting ecosystem responses to climate variability
 1126 and change. *Environmental Research Letters* 14. <https://doi.org/10.1088/1748-9326/ab33cb>.
- 1127 118. Brown, L.A., Morris, H., Leblanc, S., Bai, G., Lanconelli, C., Gobron, N., Meier, C., and
 1128 Dash, J. (2023). HemiPy: A Python module for automated estimation of forest biophysical
 1129 variables and uncertainties from digital hemispherical photographs. *Methods in Ecology*
 1130 *and Evolution* 14, 2329–2340. <https://doi.org/10.1111/2041-210X.14199>.
- 1131 119. Brown, L.A., and Leblanc, S.G. (2024). CoverPy: Automated estimates of plant area index,
 1132 vegetation cover, crown cover, crown porosity, and uncertainties from digital cover
 1133 photography in Python. *SoftwareX* 27, 101767.
 1134 <https://doi.org/10.1016/j.softx.2024.101767>.

- 1135 120. Meier, C. (2021). TOS Protocol and Procedure: DHP - Measurement of Leaf Area Index.
1136 NEON.DOC.014039. NEON (National Ecological Observatory Network).
- 1137 121. Karan, M. ed. (2015). SuperSites Vegetation Monitoring Protocols V1.21. TERN
1138 Australian SuperSite Network. <https://doi.org/10.4227/05/56134C1310F16>.
- 1139 122. Linnenbrink, J., Milà, C., Ludwig, M., and Meyer, H. (2024). kNNDM CV: *k*-fold nearest-
1140 neighbour distance matching cross-validation for map accuracy estimation. *Geosci. Model*
1141 *Dev.* 17, 5897–5912. <https://doi.org/10.5194/gmd-17-5897-2024>.
- 1142 123. Brown, L.A., Meier, C., Morris, H., Pastor-Guzman, J., Bai, G., Lerebourg, C., Gobron, N.,
1143 Lanconelli, C., Clerici, M., and Dash, J. (2020). Evaluation of global leaf area index and
1144 fraction of absorbed photosynthetically active radiation products over North America using
1145 Copernicus Ground Based Observations for Validation data. *Remote Sensing of*
1146 *Environment* 247, 111935. <https://doi.org/10.1016/j.rse.2020.111935>.
- 1147 124. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. (2017).
1148 Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of*
1149 *Environment* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- 1150 125. Šavrič, B., Patterson, T., and Jenny, B. (2019). The Equal Earth map projection.
1151 *International Journal of Geographical Information Science* 33, 454–465.
1152 <https://doi.org/10.1080/13658816.2018.1504949>.
- 1153 126. Vancutsem, C., Achard, F., Pekel, J.-F., Vieilledent, G., Carboni, S., Simonetti, D., Gallego,
1154 J., Aragão, L.E.O.C., and Nasi, R. (2021). Long-term (1990–2019) monitoring of forest
1155 cover changes in the humid tropics. *Science Advances* 7, eabe1603.
1156 <https://doi.org/10.1126/sciadv.abe1603>.
- 1157

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SI20260923compressed.pdf](#)