nature portfolio

https://doi.org/10.1038/s41586-026-11284-y

Supplementary information

Widespread deciduousness reveals diverse Amazonian phenological strategies

In the format provided by the authors and unedited

Supplementary Information for

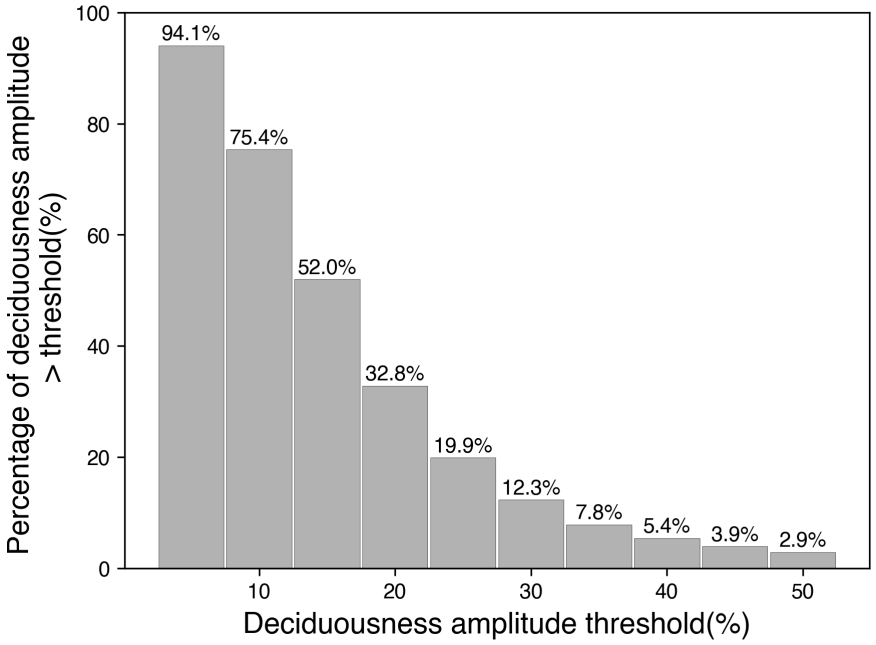
Widespread deciduousness reveals diverse Amazonian phenological strategies

Author List: Guangqin Song, Alice C. Hughes, Xu Lian, Xiang Zeng, Scott R. Saleska, Alfredo Huete, Loren P. Albert, Xiangtao Xu, Josep Penuelas, Constantin M. Zohner, Matteo Detto, S. Joseph Wright, Bruce W. Nelson, Shuli Chen, Zhengfei Guo, Yadvinder Malhi, Xing Li, Ziyu Lin, Yanjun Su, Jingfeng Xiao, Yingyi Zhao, Peng Zhu, Lingli Liu, Jin Wu

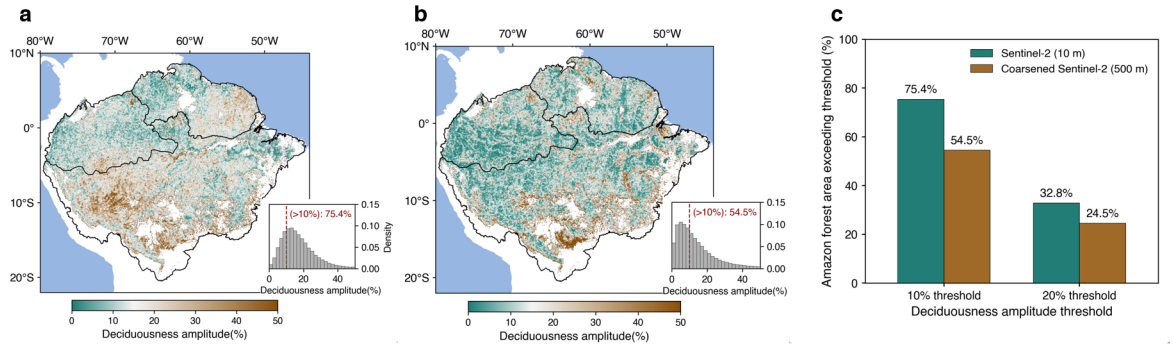
Corresponding author: jinwu@hku.hk

Contents:

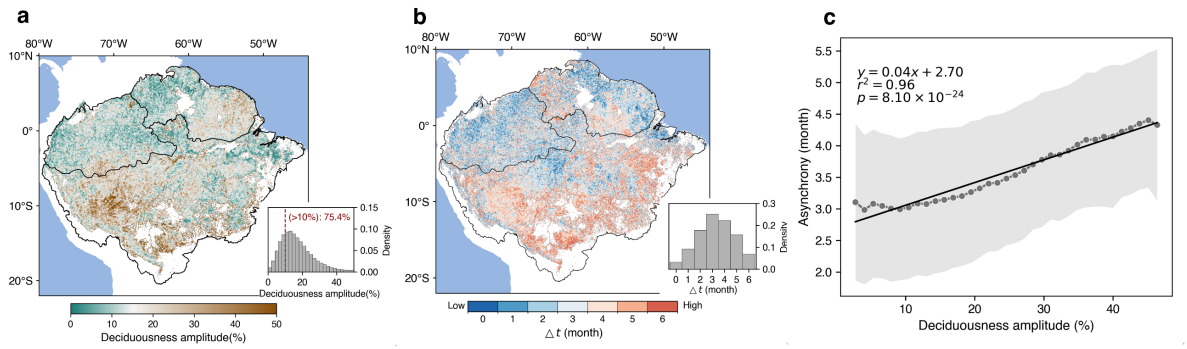
Supplementary Figs. 1 to 28
Supplementary Tables 1 to 5



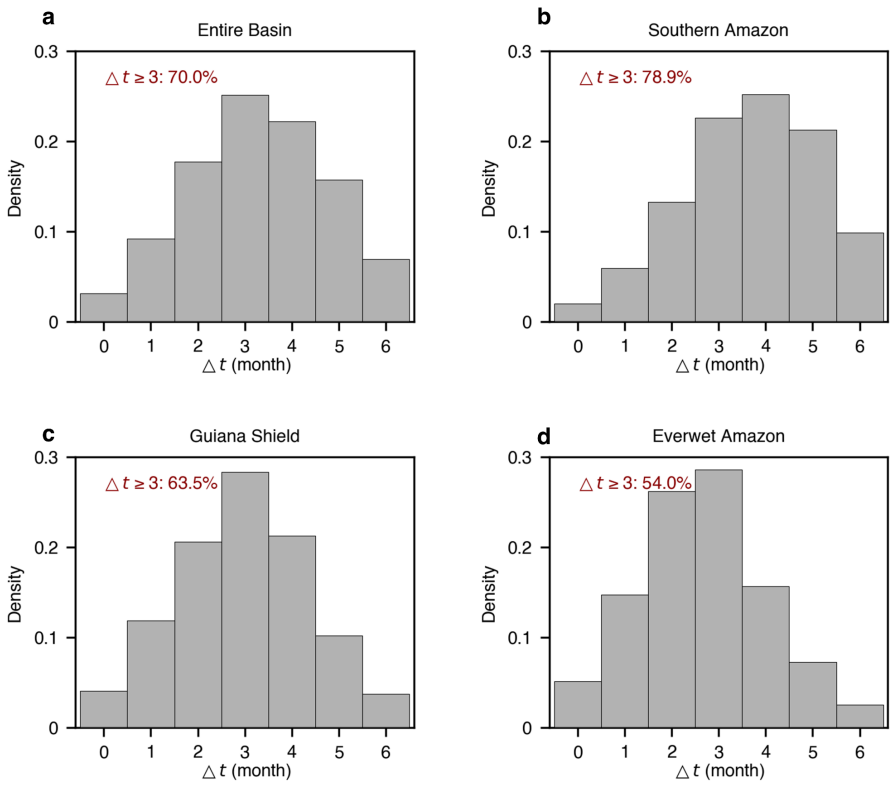
Supplementary Figure 1. Percentage of Amazon forest area exceeding different deciduousness-amplitude thresholds.



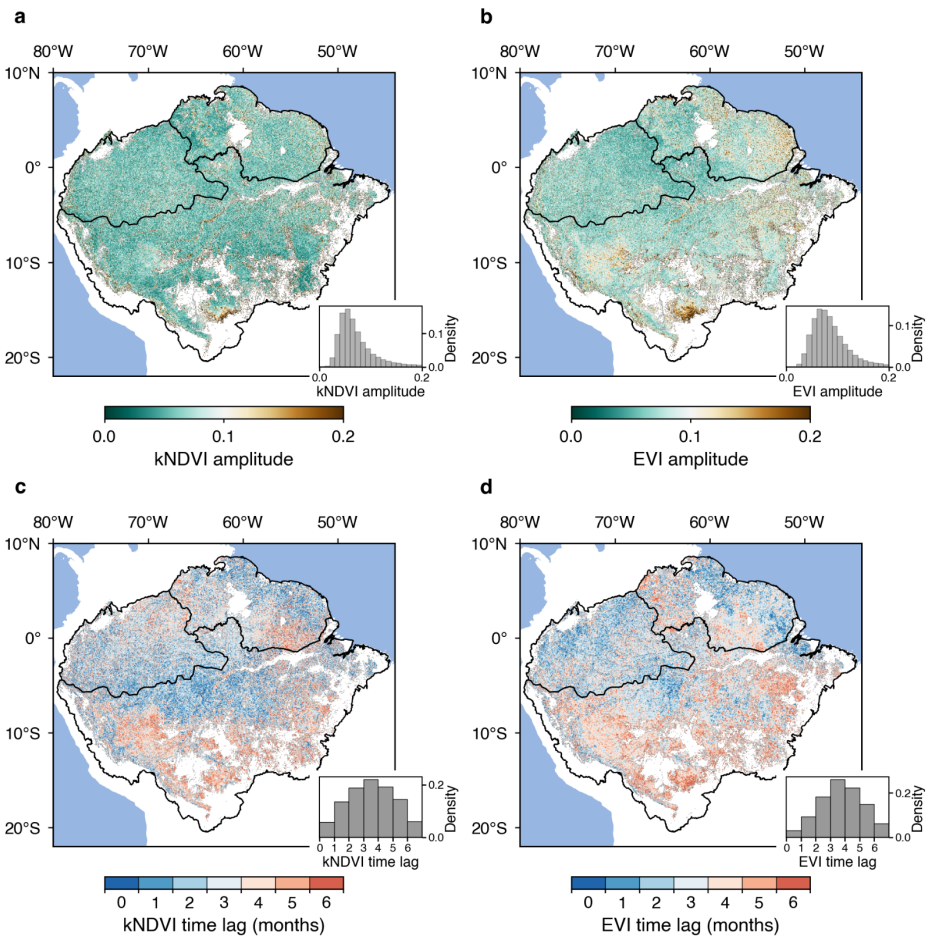
Supplementary Figure 2. Sensitivity of Amazon deciduousness estimates to spatial resolution. a, Deciduousness amplitude derived from the original 10-m Sentinel-2 observations**. b,** Deciduousness amplitude derived after aggregating the original Sentinel-2 reflectance to approximately 500 m before spectral unmixing, while retaining the grid- and year-specific endmembers derived from the 10-m imagery. **c,** Percentage of Amazon forest area exceeding deciduousness-amplitude thresholds of 10% and 20% at 10 m and 500 m resolution.



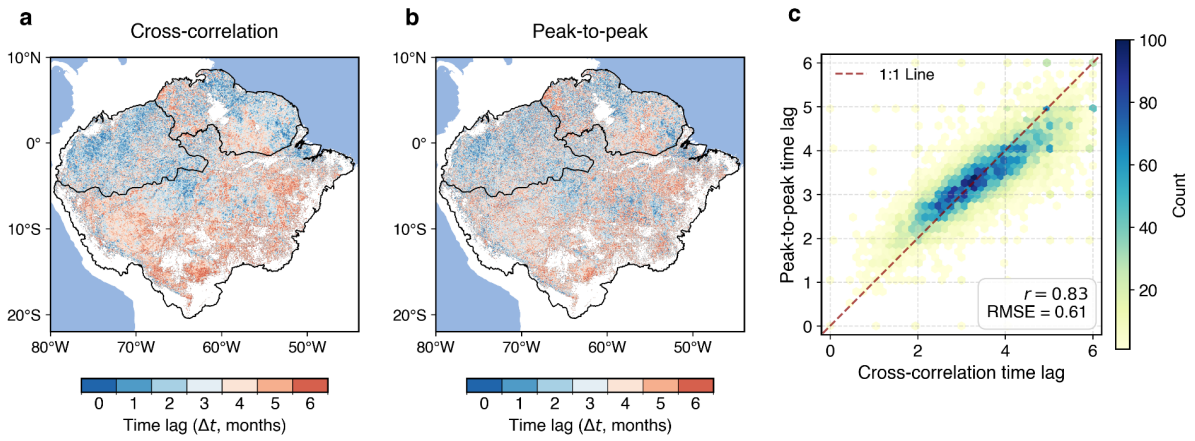
Supplementary Figure 3. Deciduousness amplitude and phenological asynchrony across the Amazon. a, Spatial distribution of deciduousness amplitude, defined as the seasonal magnitude of the leafless canopy fraction. The inset shows the basin-wide frequency distribution. **b,** Spatial distribution of phenological asynchrony (Δt). The inset shows the basin-wide frequency distribution of Δt. **c,** Binned relationship between deciduousness amplitude and phenological asynchrony across the basin. Points show the mean Δt within each deciduousness-amplitude bin, and the shaded band indicates ±1 standard deviation of pixel-level asynchrony within each bin. The black line indicates a least-squares linear regression fitted to the binned means (n = 35 bins), with statistical significance assessed using a two-sided test ($p = 8.10 \times 10^{-24}$).



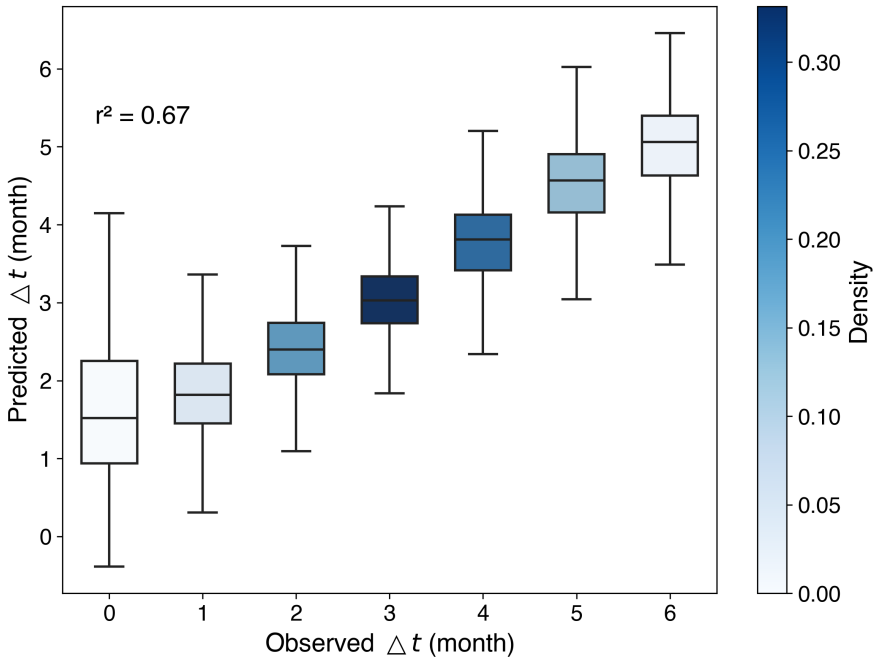
Supplementary Figure 4. Histogram distributions of phenological asynchrony (Δt) between deciduousness and greenness for the (a) entire Amazon basin, (b) Southern Amazon, (c) Guiana Shield, and (d) Everwet Amazon.



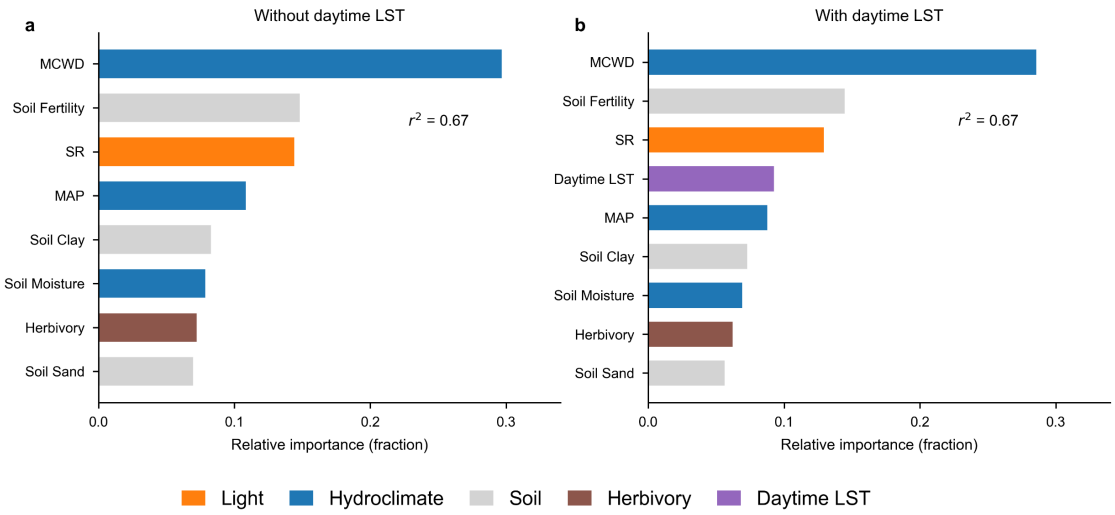
Supplementary Figure 5. Spatial variation in seasonal amplitude and deciduousness-greenness time lags for kNDVI and EVI across the Amazon basin. **a, b**, Spatial distributions of seasonal amplitude for kernel NDVI (kNDVI; **a**) and Enhanced Vegetation Index (EVI; **b**). Insets show the basin-wide frequency distributions of amplitude values. Overall, both indices exhibit similar spatial patterns, with the largest amplitudes occurring in the southwestern and central transition zones, although amplitudes are generally smaller for kNDVI than for EVI. **c, d,** Spatial distributions of the time lag (in months) between deciduousness and the seasonality of kNDVI (**c**) and EVI (**d**). Here, time lag is defined as the temporal offset that maximizes the correlation between the deciduousness time series and the corresponding greenness-index time series.



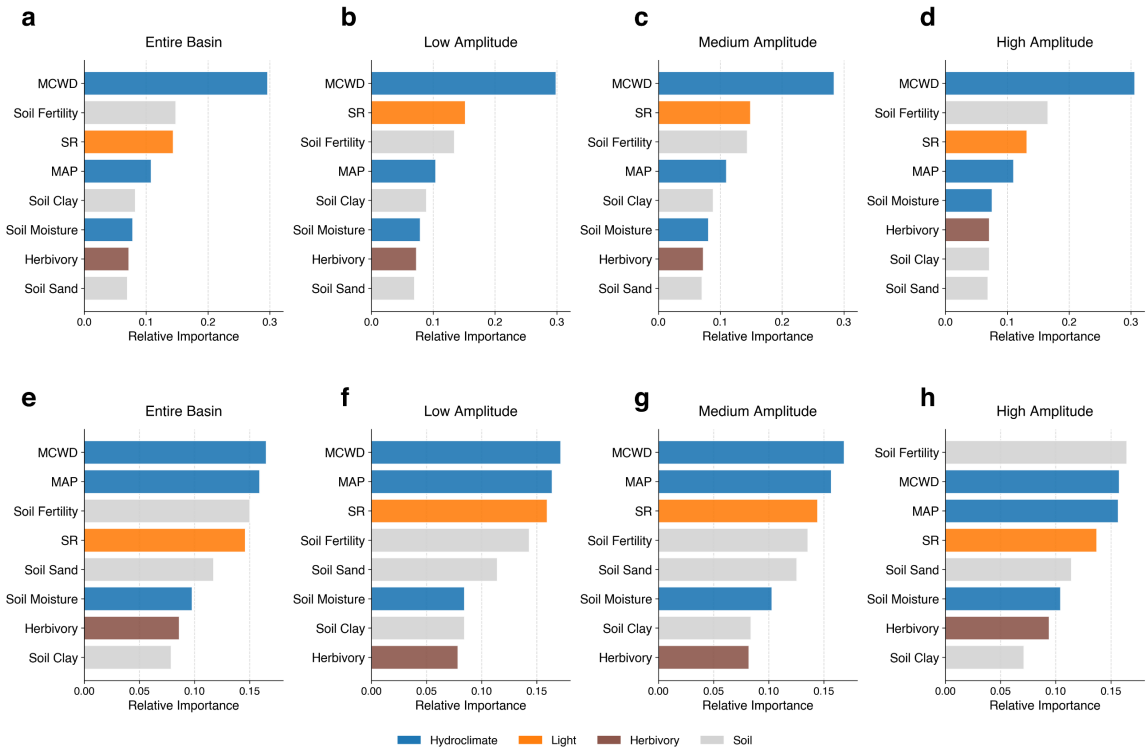
Supplementary Figure 6. Comparison of phenological time-lag estimates derived using two methods. a, Time lag (Δt) between deciduousness and greenness estimated using the maximized cross-correlation method, which incorporates the full seasonal cycle and is used in the main analysis. **b,** Time lag estimated using the peak-to-peak method, defined as the month of maximum greenness minus the month of maximum deciduousness. **c,** Regional-scale comparison of the two estimates after aggregation to 5×5-pixel blocks to reduce noise arising from integer-valued lag estimates (n = 11,251 valid blocks). The dashed red line indicates the 1:1 relationship.



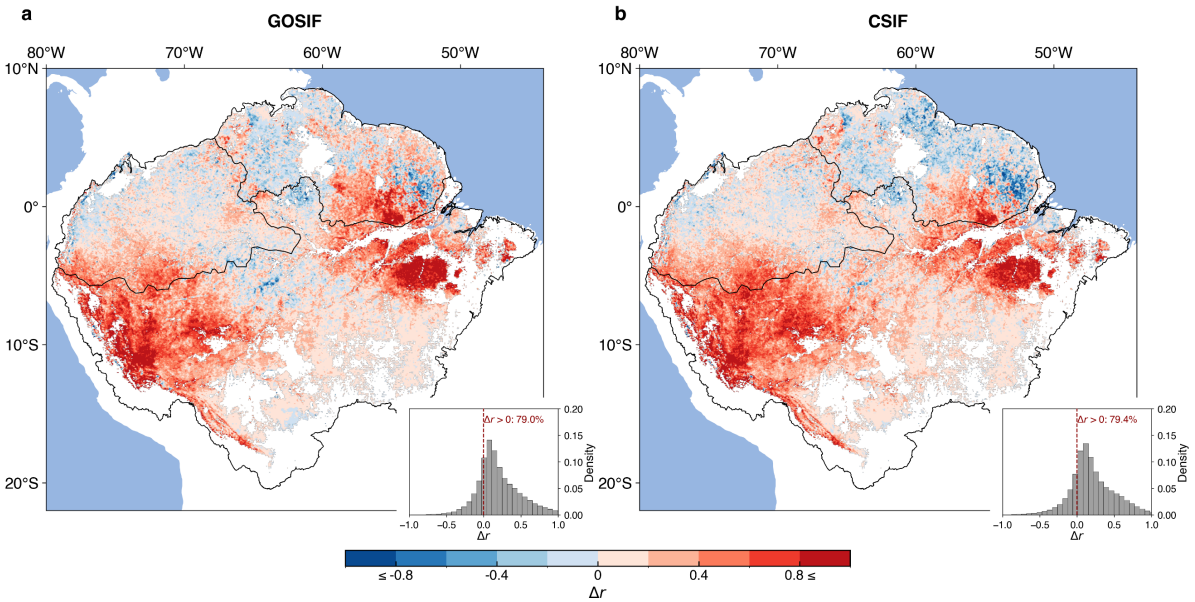
Supplementary Figure 7. Performance evaluation of the XGBoost model in predicting deciduousness-greenness asynchrony (Δ*t*) across the Amazon rainforest using 10-fold cross-validation. Boxplots show the distribution of predicted Δt values grouped by observed Δ*t* classes (0-6 months) based on pooled cross-validation predictions. The sample sizes of the observed Δ*t* classes were n = 686, 12,172, 43,642, 74,916, 57,830, 31,549 and 5,228, respectively. The center line indicates the median, the box represents the interquartile range (IQR), and the whiskers extend to 1.5 × IQR. The colour scale indicates the density of observations in each observed Δt class. Model accuracy is quantified by the squared correlation coefficient (r^2).



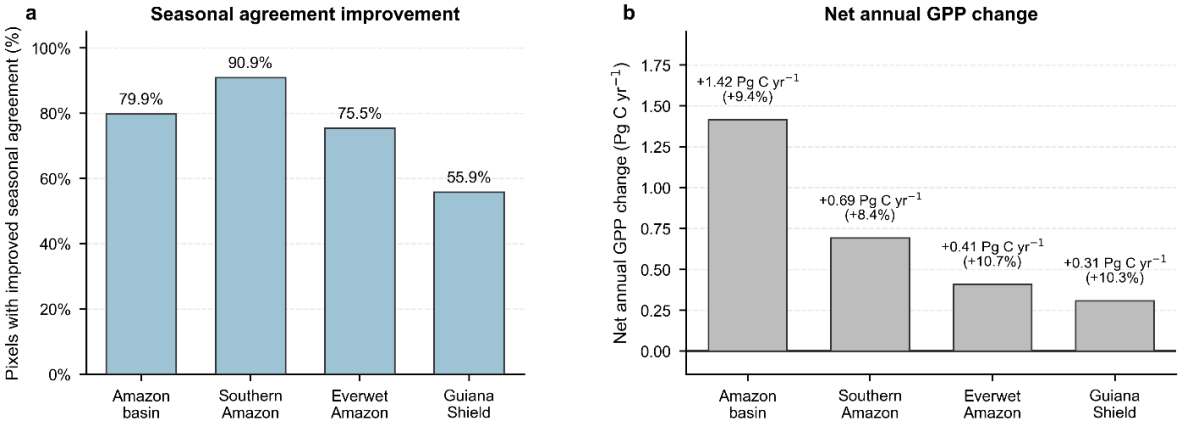
Supplementary Figure 8. Sensitivity of the environmental-driver analysis to the inclusion of daytime land-surface temperature (LST). a, Relative importance of environmental predictors in the original driver model without LST. **b,** Relative importance after adding basin-wide MODIS daytime LST averaged over 2019-2021 as an additional predictor. Predictor colors indicate the corresponding driver groups: light, hydroclimate, soil properties, herbivory, and daytime LST. MCWD, maximum cumulative water deficit; SR, solar radiation; MAP, mean annual precipitation.



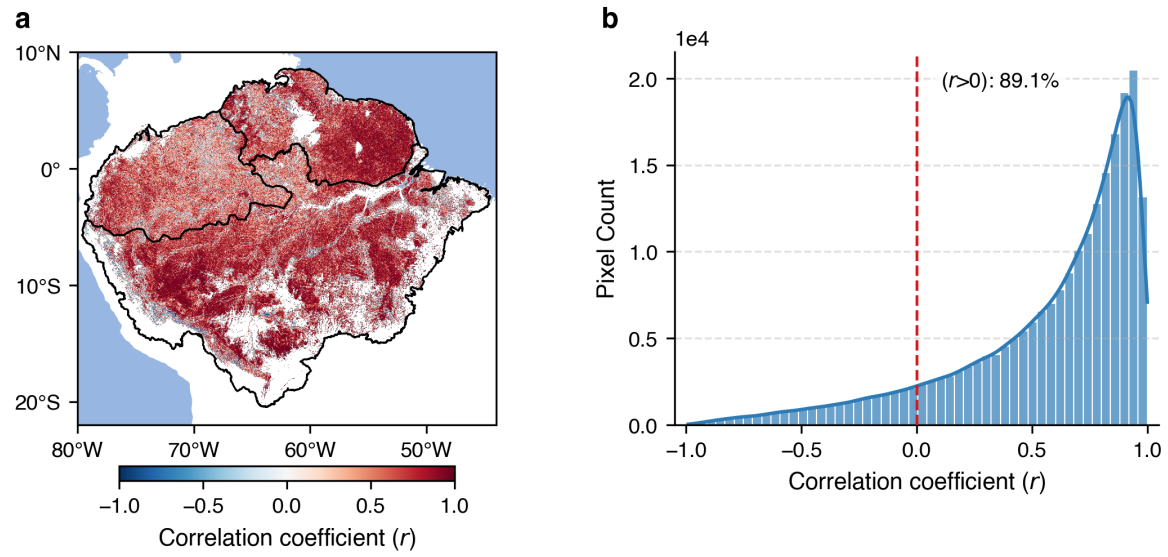
Supplementary Figure 9. Environmental drivers of phenological asynchrony and deciduousness amplitude across the deciduousness-amplitude gradient. a-d, SHAP-based relative importance of environmental predictors for phenological asynchrony across the entire Amazon basin and within low-, medium-, and high-amplitude classes. Phenological asynchrony was quantified as the time lag (Δt) between deciduousness and greenness seasonality that yielded the strongest correlation between the two metrics. **e-h**, Corresponding predictor importance for deciduousness amplitude. The XGBoost models incorporated multiple predictor dimensions, including hydroclimate (MCWD, MAP and soil moisture), light availability (SR), soil properties (soil fertility, sand and clay fractions), and herbivory pressure (the proportion of leaf area consumed by herbivores). Bars show mean absolute SHAP values, normalized within each panel, with predictors ranked by relative importance. Amplitude classes were defined by the lower, middle, and upper terciles of deciduousness amplitude, corresponding to approximately amplitude < 12%, 12% ≤ amplitude < 20% and amplitude ≥ 20% in this dataset.



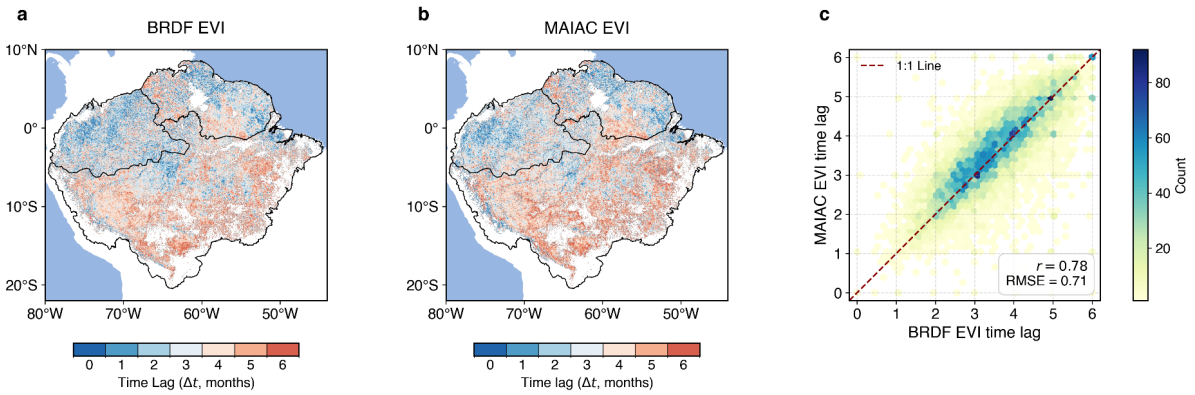
Supplementary Figure 10. Incorporating deciduousness and associated leaf age demography improves simulations of photosynthetic seasonality across the Amazon rainforest. **a, b,** Spatial patterns of model-performance gain, expressed as the change in the correlation coefficient (Δr) between simulated GPP seasonality and satellite-derived solar-induced chlorophyll fluorescence (SIF), for the LD-LUE ensemble relative to the conventional LUE ensemble (Conv-LUE). Model performance is evaluated independently against two SIF products: GOSIF (**a**) and CSIF (**b**). Insets show the basin-wide distributions of Δr, with the red dashed line marking $\Delta r = 0$ and the annotated values indicating the percentage of forest pixels with $\Delta r > 0$.



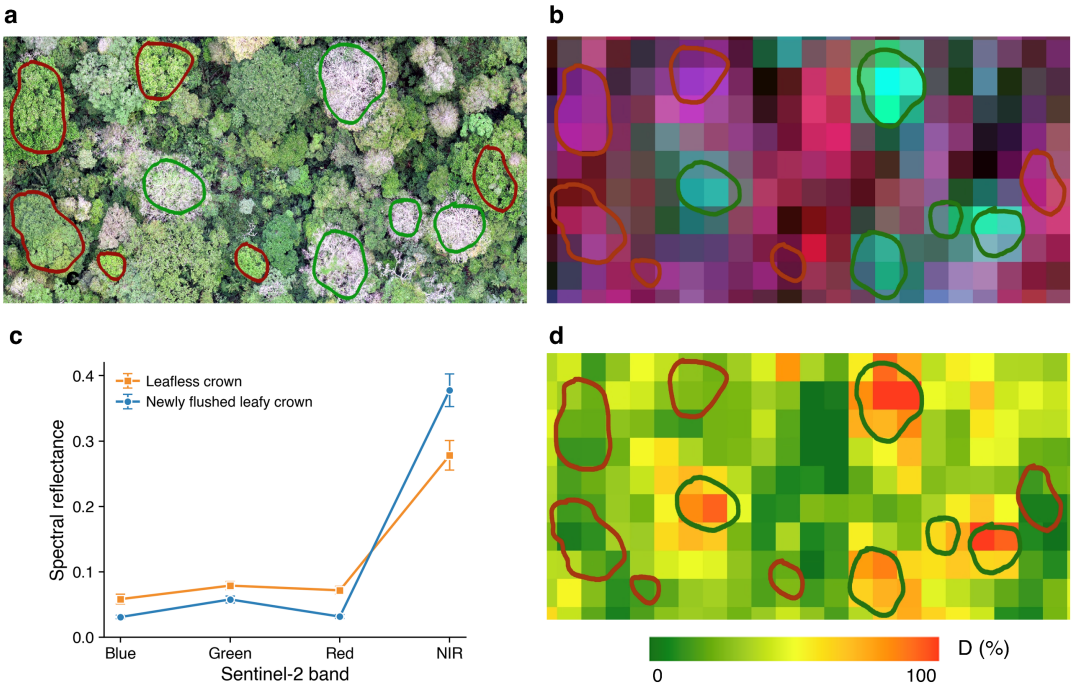
Supplementary Figure 11. Regional differences in model-simulated GPP seasonality and annual magnitude between the LD-LUE and conventional LUE ensembles. a, Fraction of forest pixels showing improved seasonal agreement with satellite SIF under the full LD-LUE ensemble relative to the conventional LUE ensemble, defined as ($\Delta r>0$), for the Amazon basin and three subregions. **b,** Net annual GPP change between the LD-LUE and conventional LUE ensembles, calculated by retaining the sign of the pixel-wise annual GPP differences and spatially integrating them over each region. Values above the bars show the net annual GPP change in Pg C yr^{-1}, with percentages in parentheses indicating the corresponding change relative to annual GPP simulated by the conventional LUE ensemble for each region. Positive values indicate higher basin- or region-integrated annual GPP under LD-LUE.



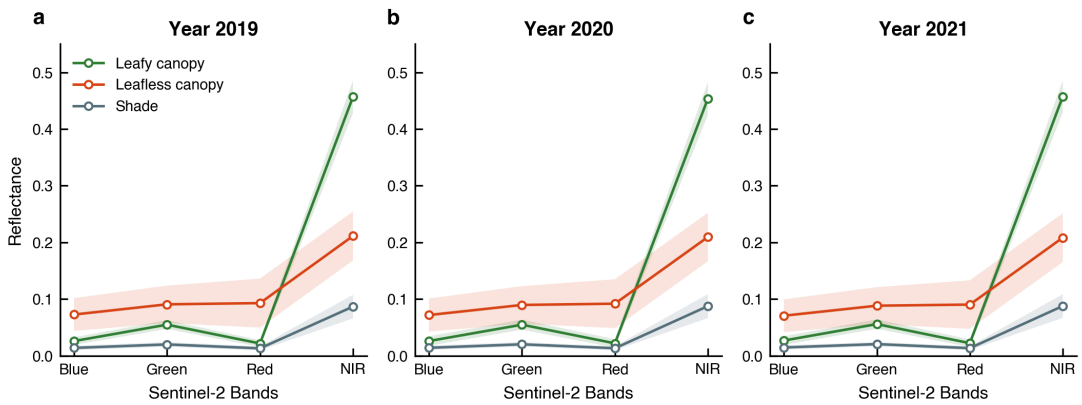
Supplementary Figure 12. Comparison of EVI seasonality derived from standard BRDF-corrected MODIS and MAIAC products across the Amazon basin. a, Pixel-wise temporal correlation (Pearson's r) between the seasonal cycles of BRDF-corrected MODIS EVI (MCD43A1) and MAIAC EVI, calculated from 12 paired monthly climatological values per pixel (n = 12). **b,** Frequency distribution of the correlation coefficients shown in (a). Overall, the two products are highly consistent, with 89.1% of pixels exhibiting positive correlations.



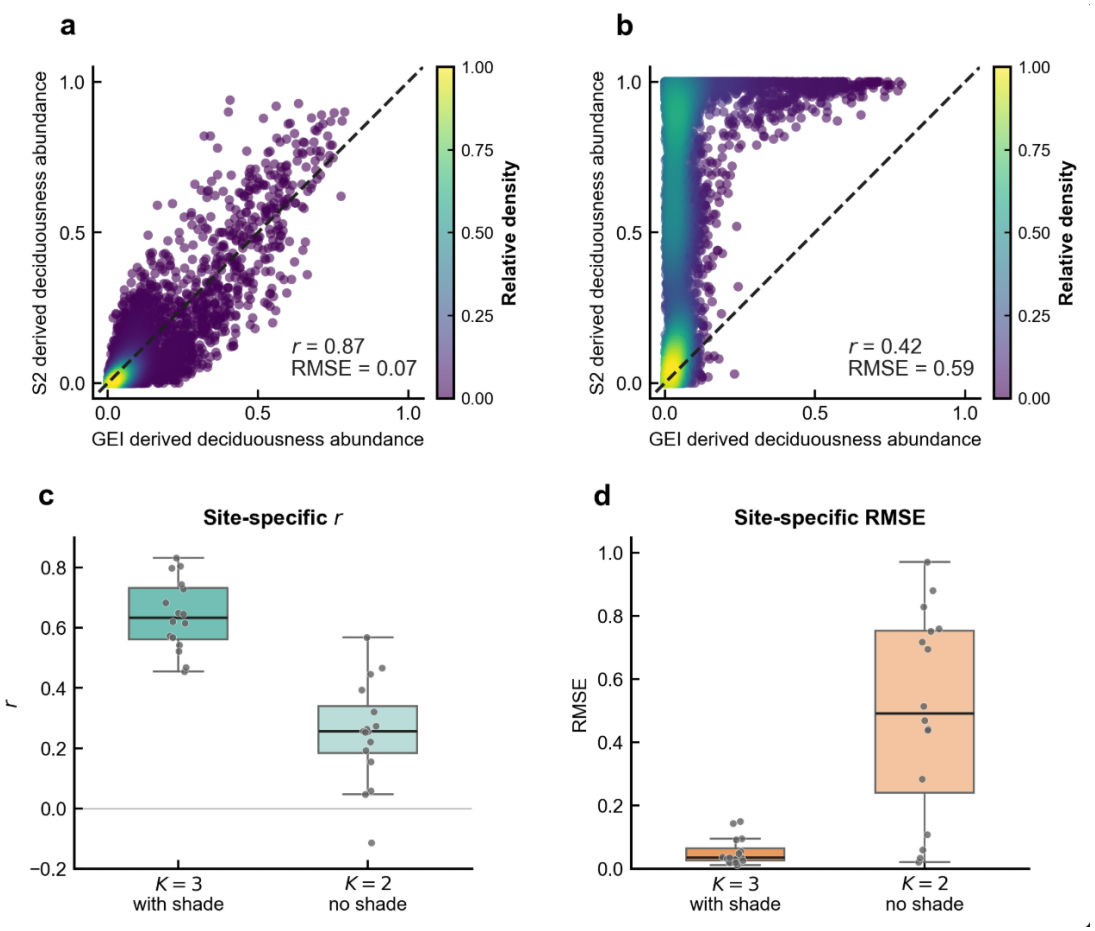
Supplementary Figure 13. Comparison of the deciduousness-greenness time lag derived from two MODIS EVI products. a, Spatial distribution of the time lag (Δ*t*, months) between deciduousness and greenness calculated using BRDF-corrected MODIS EVI. **b,** Spatial distribution of the time lag (Δ*t*, months) between deciduousness and greenness calculated using MAIAC EVI. **c**, Regional comparison of aggregated time-lag values between the two products using 5 × 5-pixel blocks (*n* = 11,204 valid blocks). This block size was used because time-lag values, expressed in months, are discrete integers and are therefore susceptible to pixel-level quantification noise. Aggregating to 5 × 5-pixel blocks reduces this error and provides a fair comparison. The dashed red line indicates the 1:1 relationship.



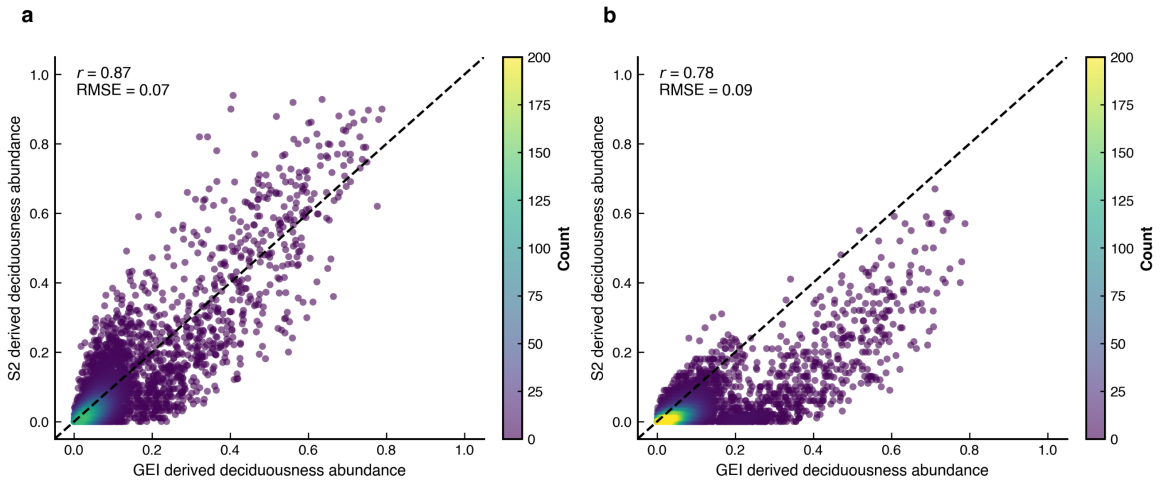
Supplementary Figure 14. Illustrative example showing the spectral distinction between a newly flushed leafy canopy and leafless crowns using co-registered drone and Sentinel-2 imagery. a, High-resolution drone image acquired on 25 March 2019, showing leafy crowns dominated by newly flushed leaves (red outlines) and crowns with exposed bare branches (green outlines). **b,** Corresponding Sentinel-2 false-color (NIR-R-G) composite acquired on 21 March 2019. **c,** Mean reflectance spectra of the two canopy endmembers across the Blue, Green, Red, and Near-Infrared (NIR) bands, demonstrating that the leafy canopy (dominated by new leaves) and leafless crowns differ distinctively in both overall brightness and multi-band spectral shape. Points show mean reflectance, and error bars indicate the standard error of the mean across sampled pixels (n = 21 pixels from newly flushed leafy crowns and n = 18 pixels from leafless crowns). **d,** Satellite-derived deciduousness fraction (D%) across the same scene, illustrating that leafless crowns are clearly distinguished from crowns undergoing new leaf flush using our knowledge-guided spectral unmixing approach.



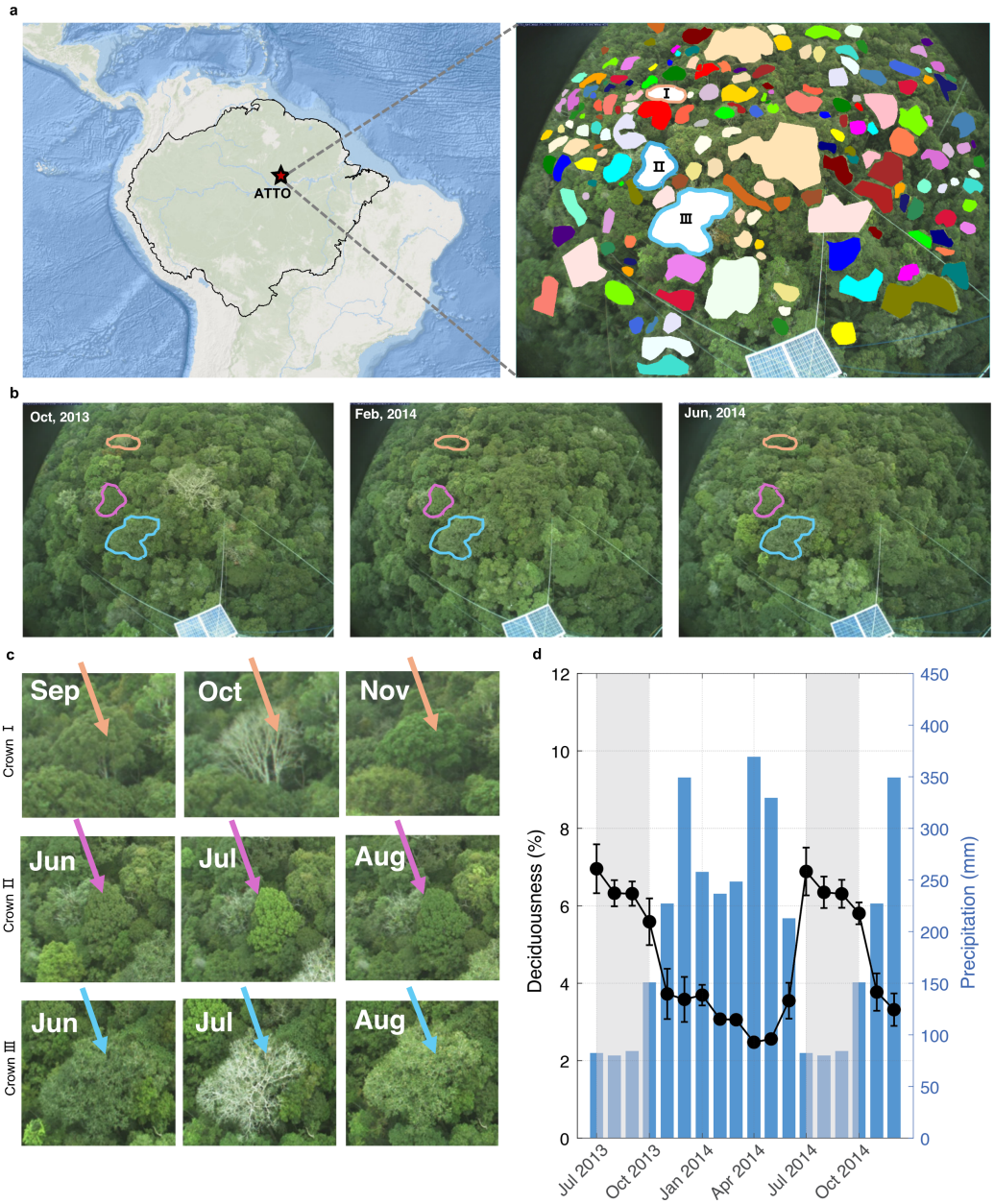
Supplementary Figure 15. Spatial variability and temporal stability of dynamically derived endmember spectra across the Amazon basin (2019-2021). a-c, Basin-wide mean spectral reflectance curves for three endmembers of leafy canopy (green), leafless canopy (orange), and shade (grey) for the years 2019 (**a**), 2020 (**b**), and 2021 (**c**), respectively. Solid lines show mean reflectance values derived using the spatially and temporally adaptive IG-ECAE method across 76,276 individual 10 km × 10 km tiles. Shaded ribbons indicate ±1 standard deviation.



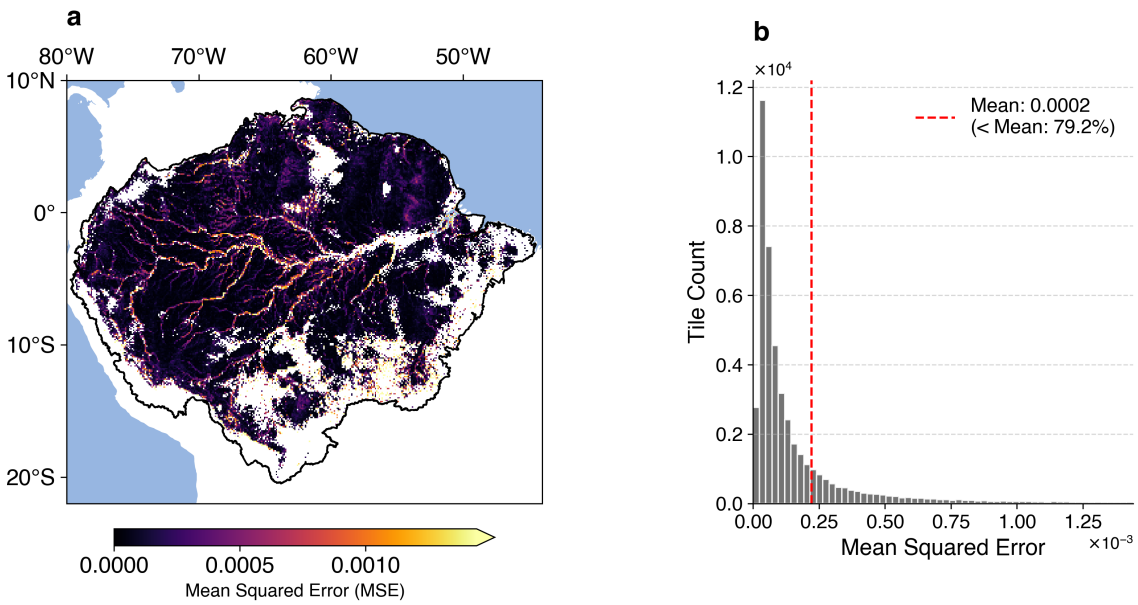
Supplementary Figure 16. Sensitivity of deciduousness retrieval to the inclusion of an explicit shade endmember. a, b, Comparison of Sentinel-2 (S2)-derived and Google Earth imagery (GEI)-derived deciduousness estimates using the original three-endmember framework with shade (*K*=3; **a**) and the independently retrained two-endmember framework without shade (*K*=2; **b**). Each point in panels a and b represents an individual validation unit (n = 6,058 patches), defined as a 100 m × 100 m sub-grid within each 2 km × 2 km GEI validation site, as shown in Extended Data Fig. 1. Colors indicate the relative density of observations, and dashed lines denote the 1:1 relationship. The three-endmember framework showed substantially stronger agreement with the independent GEI observations (r=0.87, RMSE = 0.07) than the two-endmember framework (r=0.42, RMSE = 0.59). **c, d,** Distributions of site-specific correlation coefficients (**c**) and RMSE (**d**) across the 16 GEI validation sites for the two formulations, with gray points representing individual validation sites. Boxplots show the median (center line), interquartile range (box), and whiskers extending to 1.5 × IQR.



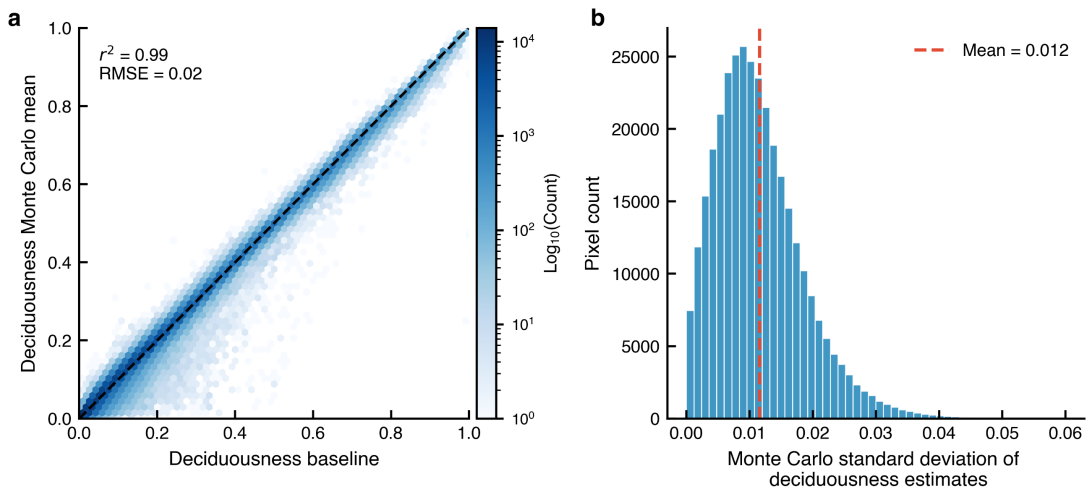
Supplementary Figure 17. Effect of shade treatment on patch-level deciduousness estimates using the three-endmember framework including shade. a, Shade fraction redistributed proportionally between leafy and leafless canopy fractions. **b,** Shade fraction retained as an independent endmember. Each point represents an individual 100 m × 100 m validation patch (n = 6,058 patches) within the 16 Google Earth imagery (GEI) validation sites. The black dashed line marks the 1:1 relationship with the benchmark, enabling direct comparison of the two approaches, using (1) benchmark deciduousness derived from image-specific supervised classification of Google Earth imagery (GEI) with a random forest (RF) approach across 16 Amazonian sites and (2) IG-ECAE estimates derived from the corresponding, closely date-matched Sentinel-2 (S2) time-series data. The 16 GEI sites used for this evaluation are shown in Extended Data Fig. 1.



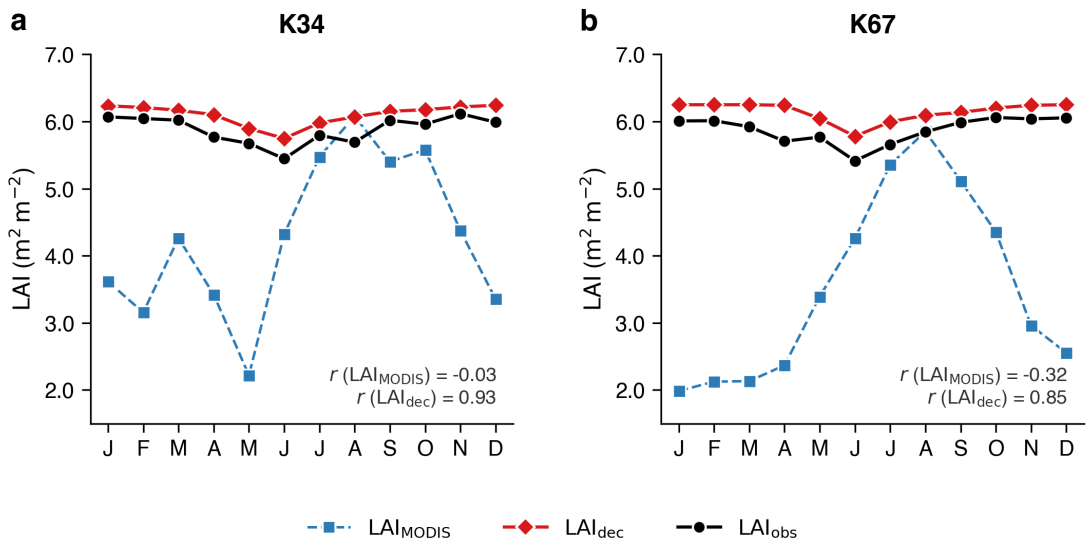
Supplementary Figure 18. Derivation of ecosystem-scale deciduousness from phenocam observations. a, Location of the Amazon Tall Tower Observatory (ATTO) site (left) and a representative phenocam image with individual tree crowns delineated (right). **b,** Seasonal changes in canopy appearance at the landscape scale across three representative months. **c,** Seasonal leaf-status dynamics of three illustrative crowns (I–III), selected for clear identification across the image sequence and shown only to illustrate the tracking procedure. **d,** Ecosystem-scale deciduousness (black line), calculated as the monthly proportion of leafless crowns among consistently tracked crowns (n = 182), together with monthly precipitation (blue bars). The line shows the mean seasonal cycle, with error bars indicating ±1 standard deviation across daily mean deciduousness within each month (n = 31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30 and 31 days for January–December, respectively). Gray shading denotes the dry season (precipitation <100 mm).



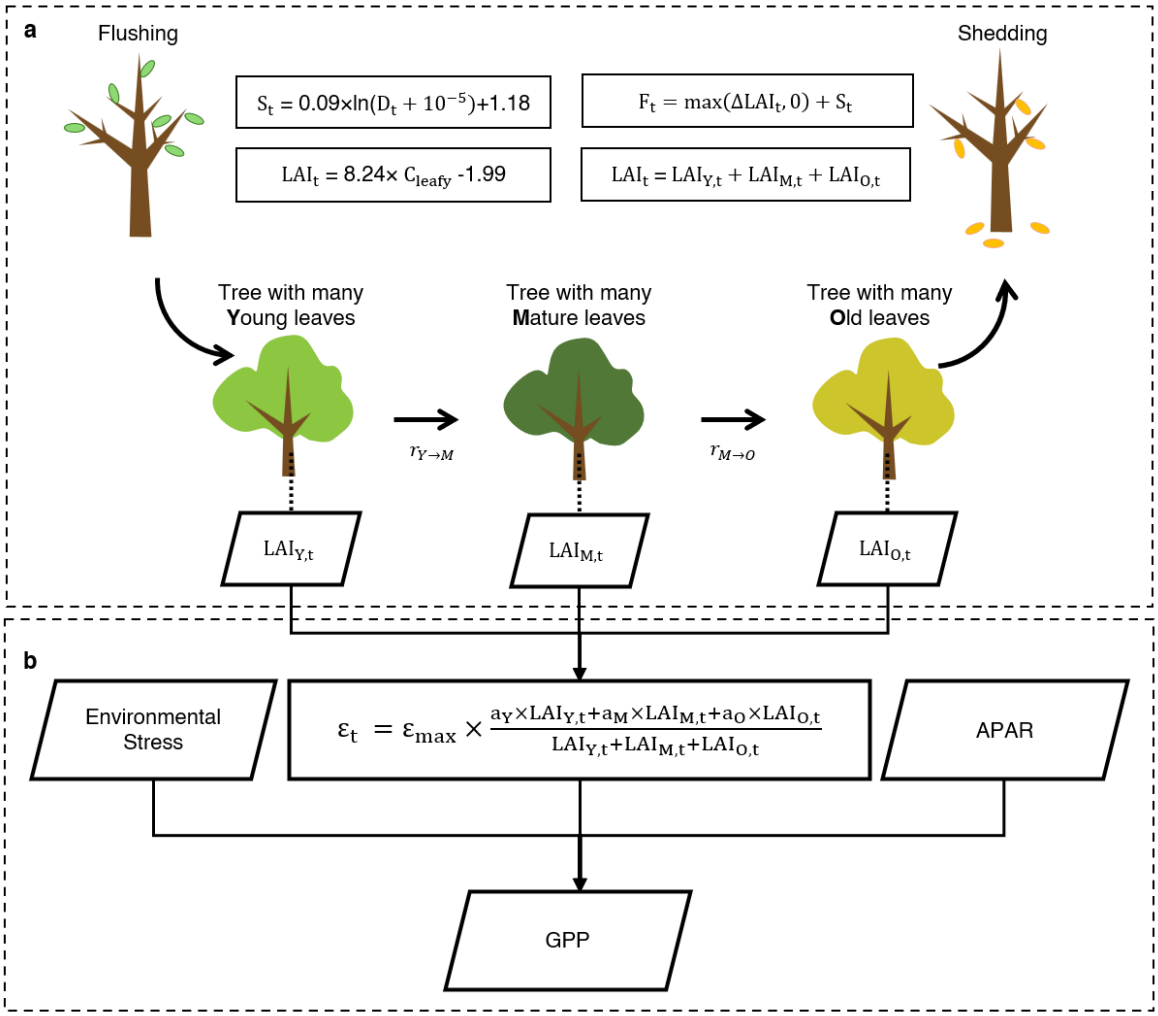
Supplementary Figure 19. Spatial pattern and distribution of spectral reconstruction error across the Amazon Basin. a, Spatial distribution of the mean squared error (MSE) between observed and reconstructed spectra per 10-km tile across the Amazon Basin. Reconstruction errors are generally low, with relatively higher MSE values (yellow/orange) occurring mainly along river corridors and water boundaries. **b,** Frequency distribution of tile-level MSE values across n = 46,038 evaluated 10-km tiles. The distribution is strongly skewed toward zero, with a mean MSE of 0.0002 (red dashed line).



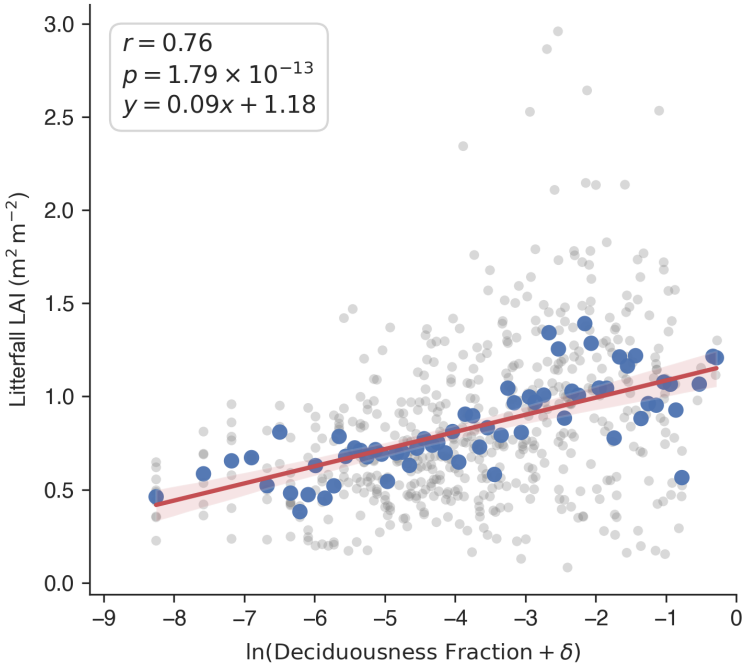
Supplementary Figure 20. Uncertainty and robustness of the spectral unmixing model assessed by Monte Carlo perturbation. a, Scatter plot comparing the no-noise baseline deciduousness fractions with the mean deciduousness fractions derived from Monte Carlo iterations (N = 50) with band-specific added Gaussian spectral noise. The comparison was performed across *n* = 334,882 evaluated pixels. The color gradient represents the pixel count on a logarithmic scale (Log_{10}) within each hexagonal bin. The black dashed line indicates the 1:1 relationship. **b**, Frequency distribution of the Monte Carlo standard deviations of deciduousness estimates across *n* = 334,882 evaluated pixels, representing the pixel-level estimated uncertainty of the unmixing results. The vertical dashed red line represents the mean standard deviation across all evaluated pixels.



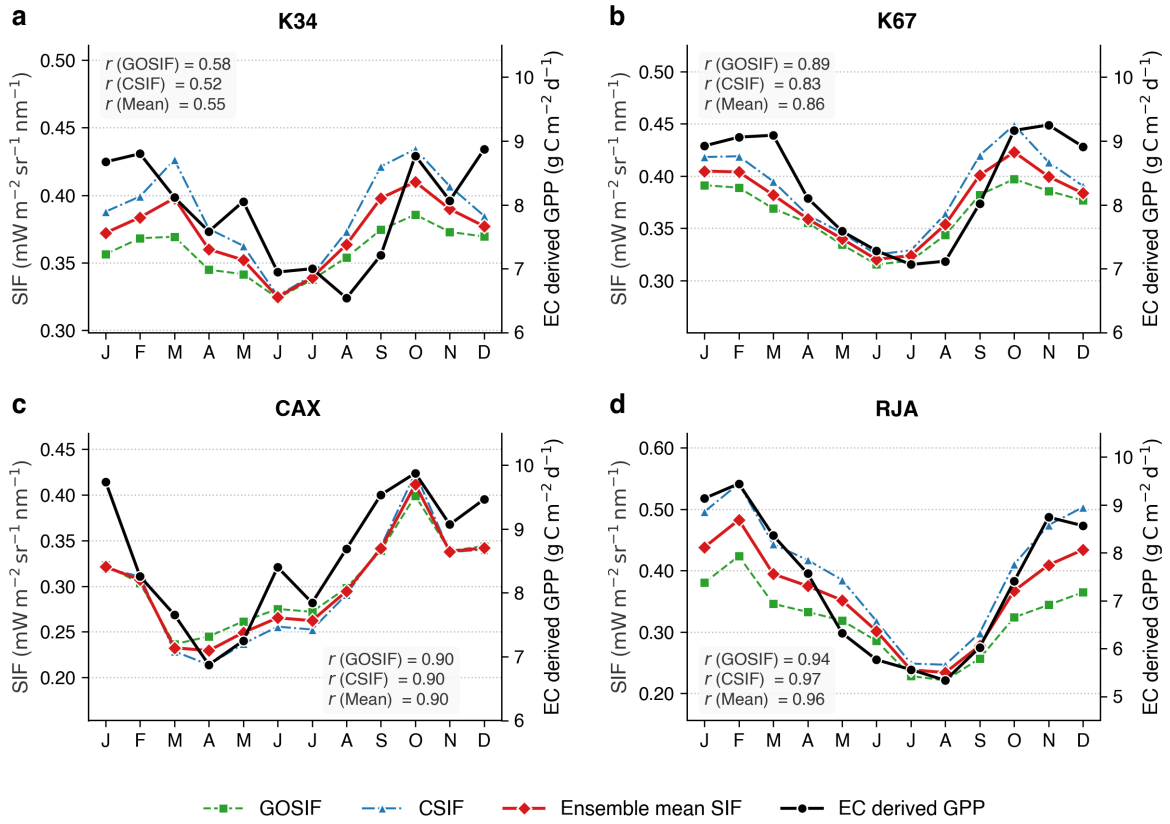
Supplementary Figure 21. Validation of the leaf demography-derived LAI (LAI_{dec}) seasonality against in-situ and satellite observations. a, b, Monthly climatology of LAI (m^2 m^{-2}) at two Amazonian evergreen forest sites: K34 (**a**) and K67 (**b**). The panels compare the seasonal trajectories of the deciduousness-driven baseline (LAI_{dec}; red solid line with diamonds), the standard satellite-derived product (LAI_{MODIS}; blue dashed line with squares), and in-situ ground measurements (LAI_{obs}; black solid line with circles). Correlation coefficients between the simulated and reference LAI seasonal cycles were calculated from 12 paired monthly climatological values (n = 12).



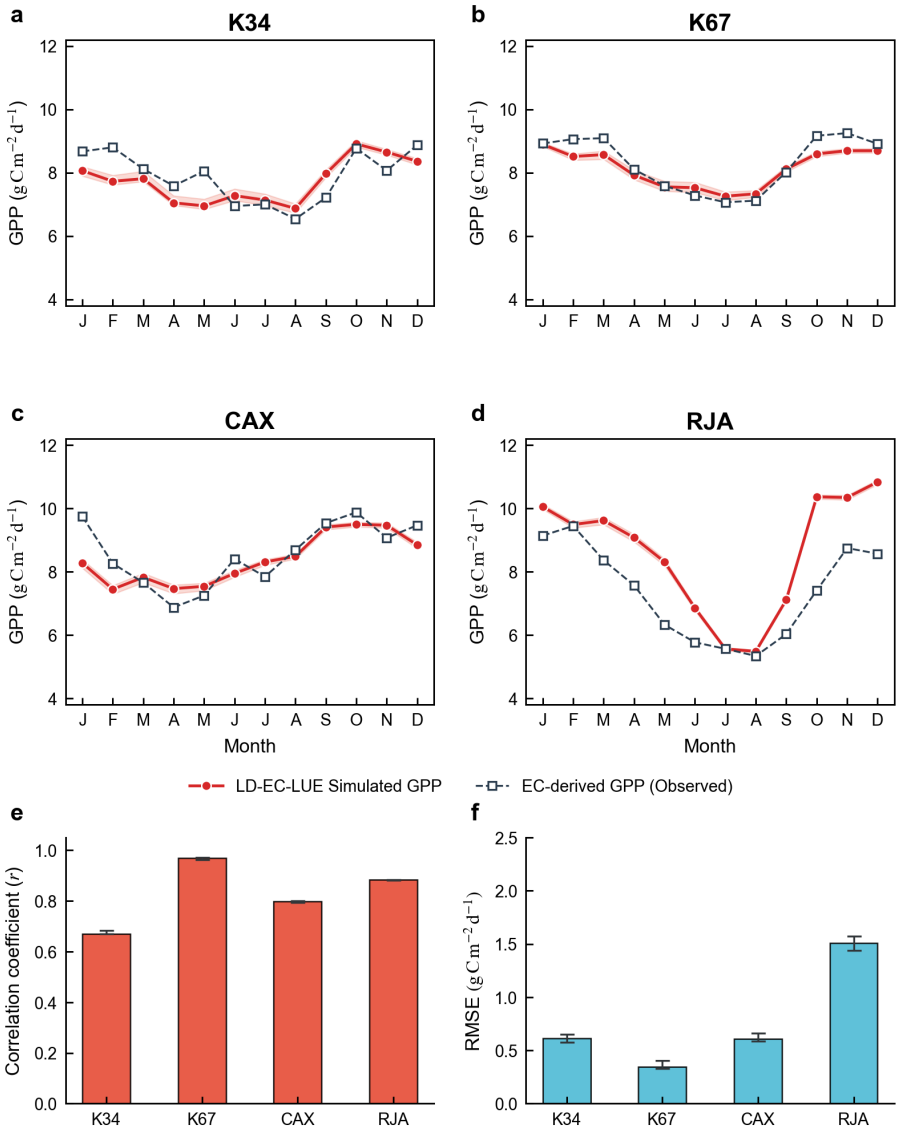
Supplementary Figure 22. Flowchart of the LD-LUE model framework. a, Schematic illustrating the derivation of leaf age dynamics, where total leaf area index (LAI) is partitioned into three age classes: young (LAI_Y), mature (LAI_M), and old (LAI_O). LAI is derived from its empirical relationship with the Sentinel-2-derived green leafy canopy fraction (C_{leafy}), as described by Eq. 9. Litterfall LAI (Shedding [S]) is estimated as a function of the deciduousness fraction (D_t), using the empirical relationship shown in Supplementary Figure 23 (Eq. 13), while flush LAI (Flushing [F_t]) is calculated as the sum of the monthly increase in LAI_t (non-negative values, constrained by a max function) and monthly litterfall (Eq. 14). Using the derived time series of Flush LAI and Litterfall LAI, along with two constants ($r_{Y \to M}$ and $r_{M \to O}$) representing the monthly fractions of leaves transitioning from the young to mature age class and from the mature to old age class, respectively, we generate time series for $LAI_{Y,t}$, $LAI_{M,t}$, and $LAI_{O,t}$ (see Eqs. 9-16). **b**, Integration of leaf age demography time series with age-dependent photosynthetic capacity to characterize seasonal variation in light use efficiency (ε_t), which is ultimately used in the LD-LUE model (together with environmental stress and APAR) to estimate the seasonal dynamics of gross primary productivity (GPP).



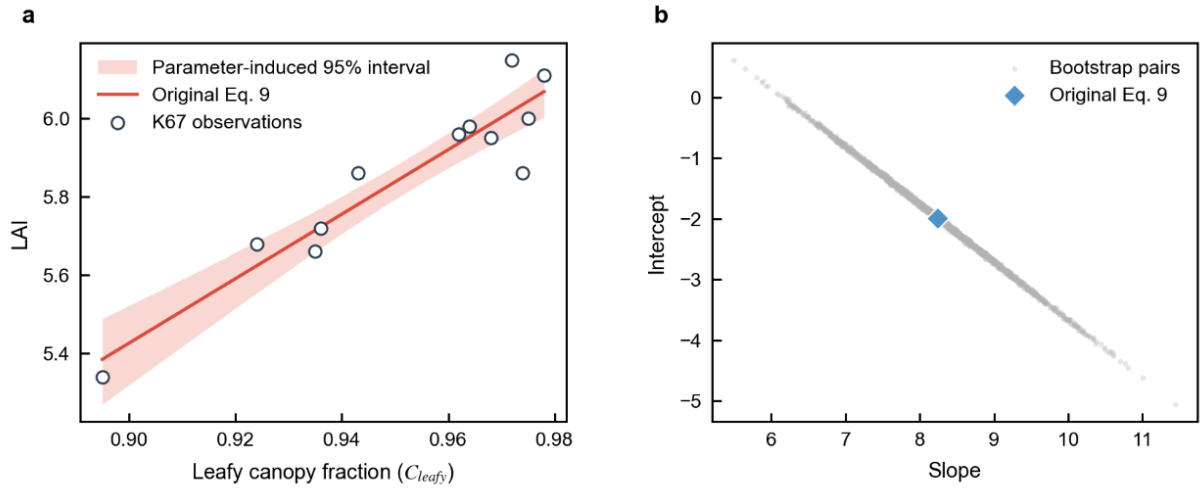
Supplementary Figure 23. Relationship between deciduousness and litterfall LAI. Litterfall LAI was calculated by converting mass-based litterfall data using the specific leaf area (SLA). Grey circles represent individual raw observations. The data were binned using a constant interval of 0.1 on a log-transformed scale (i.e., ln (Deciduousness Fraction+ **δ**), where **δ** = 10^{-5} is a small constant added before log transformation to avoid undefined values near zero deciduousness). Blue circles denote binned means (n = 65 bins). The red solid line represents the least-squares linear regression fitted to the binned means, and the light red shaded area indicates the 95% confidence interval of the regression fit. The regression significance was assessed using a two-sided test ($p = 1.79 \times 10^{-13}$).



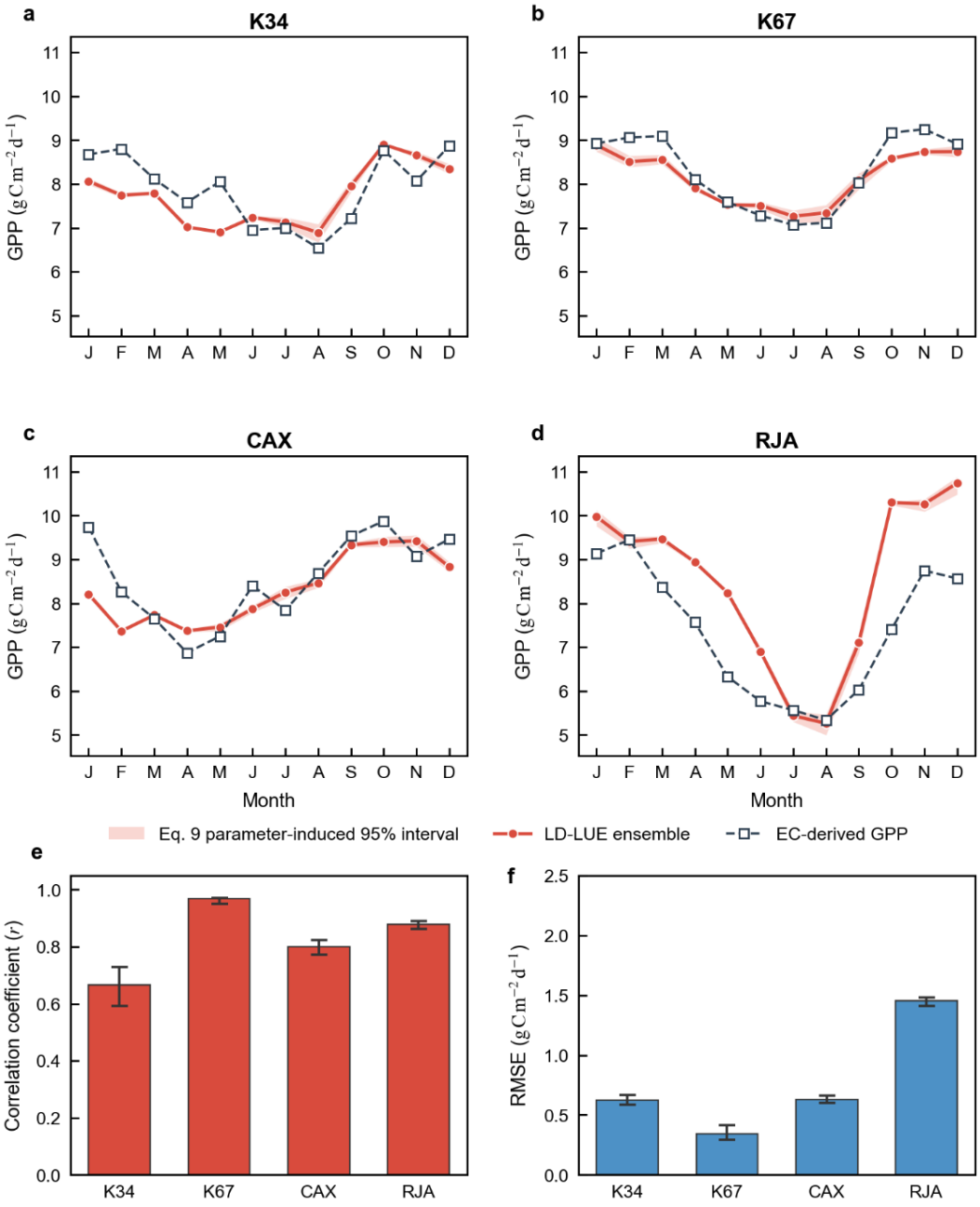
Supplementary Figure 24. Site-level seasonal correspondence between satellite-derived solar-induced chlorophyll fluorescence (SIF) and eddy covariance (EC)-derived GPP. a–d, Monthly climatologies of SIF and EC-derived GPP at four Amazonian flux-tower sites: K34 (January 2000-September 2006; a), K67 (January 2002-December 2005; b), CAX (January 2000-July 2003; c), and RJA (January 2000-October 2002; d). For each site, GOSIF and CSIF were averaged using the same year–month observations that contributed to the corresponding EC-GPP climatology. The left y-axis shows SIF (mW m^{-2} sr^{-1} nm^{-1}) from GOSIF (green dashed line with squares), CSIF (blue dash-dotted line with triangles), and their ensemble mean (red solid line with diamonds). The right y-axis shows EC-derived GPP (g C m^{-2} d^{-1}; black solid line with circles). Text within each panel reports the correlation coefficients (*r*) between EC-derived GPP and GOSIF, CSIF, and the SIF ensemble mean. Correlation coefficients were calculated from 12 paired monthly climatological values (*n* = 12) at each site.



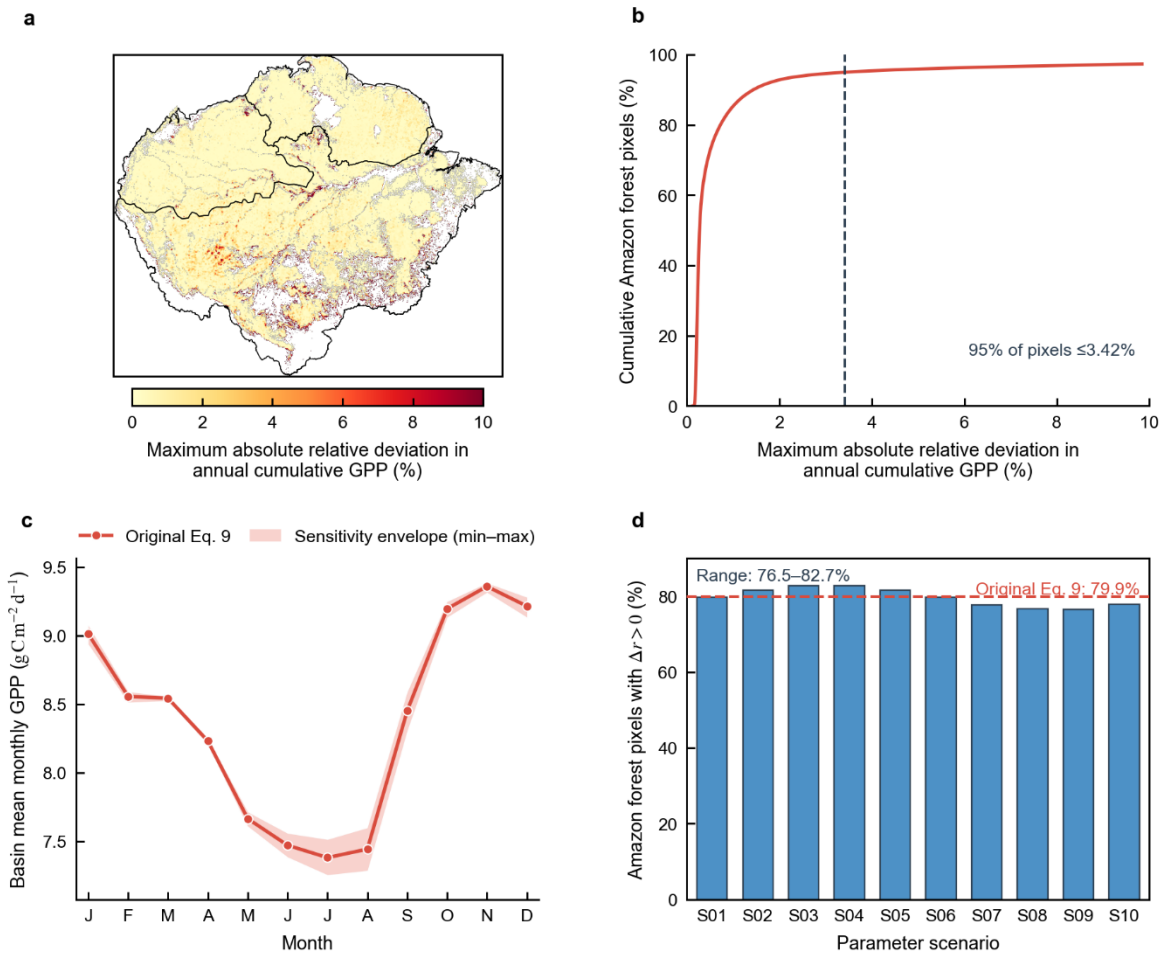
Supplementary Figure 25. Propagation of satellite-derived deciduousness uncertainty into GPP simulation using the LD-EC-LUE model at four Amazonian flux-tower sites. **a-d,** Seasonal dynamics of monthly GPP simulated by the LD-EC-LUE model (red solid lines with filled circles) compared with EC-derived GPP (black dashed lines with open squares) at sites K34 (**a**), K67 (**b**), CAX (**c**), and RJA (**d**). Red shaded regions denote the ±1 standard deviation uncertainty bounds, reflecting the nonlinear propagation of uncertainty in satellite-derived deciduousness estimates, using a deciduousness-estimate standard deviation of 0.012 derived from the Monte Carlo perturbation analysis of the IG-ECAE retrieval framework. **e,** Correlation coefficient (*r*) and **f,** Root Mean Square Error (RMSE) between simulated and EC-derived GPP across the four sites. Error bars indicate the minimum-to-maximum range of performance metrics obtained from simulations using the upper and lower bounds of the deciduousness perturbation scenarios.



Supplementary Figure 26. Bootstrap-based uncertainty in the parameters of Eq. 9. a, Relationship between leafy canopy fraction (C_{leafy}) and LAI at K67. Open circles denote the 12 paired observations used to calibrate Eq. 9, the red line shows the original fitted relationship, and the shaded band represents the parameter-induced 95% interval derived from 10,000 residual-bootstrap realizations. **b,** Joint distribution of the slope and intercept obtained from the 10,000 bootstrap realizations. Gray points represent paired bootstrap estimates, preserving the covariance between the two fitted parameters, and the blue diamond indicates the parameter combination of the original Eq. 9.



Supplementary Figure 27. Propagation of Eq. 9 parameter uncertainty into site-level GPP simulations. a–d, Monthly GPP simulations at K34 (**a**), K67 (**b**), CAX (**c**), and RJA (**d**). Red lines show the LD-LUE simulations using the original Eq. 9 parameterization, and shaded bands indicate the 95% uncertainty intervals obtained by propagating 10,000 residual-bootstrap realizations of the Eq. 9 slope–intercept parameters through the GPP modelling framework. Dashed lines with open squares show EC-derived GPP. **e, f,** Correlation coefficients (**e**) and RMSE (**f**) between simulated and EC-derived monthly GPP at the four sites. Bars show the values obtained with the original Eq. 9 parameterization, and error bars indicate the corresponding 95% bootstrap intervals arising from Eq. 9 parameter uncertainty.



Supplementary Figure 28. Basin-scale propagation of uncertainty in the parameters of Eq. 9. a, Maximum absolute relative deviation in annual GPP across the 10 representative parameterizations, relative to the original Eq. 9. **b,** Pixel-count-based cumulative distribution of the maximum absolute relative deviation, with 95% of Amazon forest pixels showing deviations of ⩽3.42%. **c,** Basin-mean monthly GPP simulated using the original Eq. 9 and the minimum–maximum range across the 10 representative parameterizations. **d,** Percentage of Amazon forest area in which LD-LUE improved seasonal agreement relative to conventional LUE ($\Delta r > 0$) under each parameterization. The dashed line indicates the result obtained using the original Eq. 9.

Supplementary Table 1. Detailed information on the site-level data used in this study, including site ID, location, data type, and time range. a, Overview of validation and model-evaluation datasets and **b**, site-level metadata for eddy-covariance and phenocam sites.

(a)

Dataset	Observation/data type	No. of records	Latitude range (°)	Longitude range (°)	Time range	Data source/reference
Eddy-covariance flux towers	EC-derived GPP time series	4 sites	-10.08 to -1.72	-61.93 to -51.46	2000-01 to 2006-09	Ref. 97
Phenocam observations	Tower-mounted RGB imagery and crown leaf-status records	5 sites	-10.08 to 9.15	-79.85 to -54.97	2010-01 to 2018-12	Ref. 28
Litterfall observations	Monthly leaf litterfall records compiled across Amazon forest sites	47 sites	-16.02 to 5.27	-73.43 to -51.41	2000-01 to 2019-03	Refs. 71, 72
Forest inventory plots	Plot-level deciduous-species proportion matched with satellite deciduousness	289 plots	-18.77 to 5.62	-79.18 to -48.29	NA	Refs. 74, 75
High-resolution Google Earth imagery	Very high-resolution 2-km × 2-km image scenes	16 scenes	-12.47 to 3.58	-75.05 to -52.49	2020-06 to 2021-09	Data collected in this study

(b)

Site ID	Site name	Latitude (°)	Longitude (°)	EC time range	Phenocam time range
01-BCI	Barro Colorado Island	9.15	-79.85	NA	2013-03 to 2017-05
02-CAX	Caxiuanã National Forest	-1.72	-51.46	2000-01 to 2003-07	NA
03-ATTO	Amazon Tall Tower Observatory	-2.14	-59.00	NA	2013-07 to 2014-12
04-K34	Manaus km 34	-2.61	-60.21	2000-01 to 2006-09	2011-12 to 2013-11
05-K67	Tapajós km 67	-2.85	-54.97	2002-01 to 2005-12	2010-01 to 2011-12
06-RJA	Reserva Jaru forest site	-10.08	-61.93	2000-01 to 2002-10	2017-12 to 2018-12

Abbreviations: EC, eddy covariance

Supplementary Table 2. Summary of the independent satellite-derived products used for comparison with Sentinel-2 deciduousness amplitude. Spatial and temporal characteristics, along with data access links, are provided for each dataset. All higher-resolution products (MCD12C1, MCD43A1 and MOD44B) were spatially aggregated to a unified 5-km grid, whereas the coarser resolution product (VOD) was resampled to match the 5-km baseline analysis configuration across the Amazon basin.

Product	Spatial Resolution	Temporal Resolution	Data Availability / URL
Evergreen forest fractional cover (MCD12C1)	0.05°	Annual	https://www.earthdata.nasa.gov/data/catalog/lpcloud-mcd12c1-061
Percent tree cover (MOD44B)	250 m	Annual	https://www.earthdata.nasa.gov/data/catalog/lpcloud-mod44b-061
EVI amplitude (MCD43A1)	500 m	Daily	https://developers.google.com/earth-engine/datasets/catalog/MODIS_061_MCD43A1
Vegetation optical depth (VOD)	25 km	Monthly	https://doi.org/10.48436/t74ty-tcx62

Supplementary Table 3. Ancillary datasets used to investigate the drivers of the spatial variation in deciduousness-greenness asynchrony across the Amazon rainforest. The table lists the eight primary predictors used in the main driver analysis and the additional land surface temperature (LST) variable included only in the sensitivity analysis.

Variable type	Variable name	Spatial resolution	Temporal resolution	Coverage period	Reference
Hydroclimate	Maximum cumulative water deficit (MCWD)	0.1°	Monthly	1981-present	Ref. 93
	Precipitation	0.1°	Monthly	1981-present	Ref. 93
	Soil moisture	0.1°	Daily	1982-2022	Ref. 98
Light	Surface downwelling shortwave radiation flux (SR)	0.1°	Monthly	1981-present	Ref. 93
Soil	Soil texture (soil clay and sand fraction)	250 m	NA	NA	Ref. 84
	Soil fertility (exchangeable base cations (Ca^{2+}, Mg^{2+}, K^{+}))	0.1°	NA	NA	Ref. 85
Herbivory	Herbivory pressure	0.08°	NA	NA	Ref. 47
Temperature	Daytime land surface temperature (LST)	1 km	8-day	2000-present	Ref. 87

Note: MCWD is defined as the most negative value of the cumulative water deficit (CWD) within a year, serving as an indicator of hydroclimatic water stress. CWD is calculated annually by subtracting a fixed evapotranspiration value (100 mm per month) from monthly precipitation. For detailed data access, see the data availability statement.

Supplementary Table 4. Baseline LUE formulations, canopy inputs and core ecophysiological processes used to generate GPP across the Amazon rainforest. EC-LUE was implemented with four alternative canopy inputs (NDVI, kNDVI, EVI and LAI), whereas the two-leaf EC-LUE and MODIS-LUE formulations used LAI. ε_{max} denotes maximum light-use efficiency. ε_{msu} and $APAR_{su}$ denote the maximum light use efficiency and absorbed photosynthetically active radiation (APAR) of sunlit leaves, respectively; ε_{msh} and $APAR_{sh}$ refer to the maximum LUE and APAR of shaded leaves. In the MODIS model, f(Ta) and f(VPD) are scalars that account for the downregulation of LUE due to temperature and vapor pressure deficit (VPD), respectively.

Formulation	Canopy input	Ecophysiological processes captured by model
EC-LUE	NDVI kNDVI EVI LAI	$GPP = PAR \times FPAR \times \varepsilon_{max} \times \min(T_s, W_s)$
Two-leaf EC-LUE	LAI	$GPP = (\varepsilon_{msu} \times APAR_{su} + \varepsilon_{msh} \times APAR_{sh}) \times \min(T_s, W_s)$
MODIS-LUE	LAI	$GPP = PAR \times FPAR \times \varepsilon_{max} \times f(T_a) \times f(VPD)$

Supplementary Table 5. Calibrated parameter estimates for the leaf-demography-driven LUE configurations and their corresponding validation performance. Optimized parameters were estimated and their uncertainty was characterized, using Markov Chain Monte Carlo (MCMC) sampling. Values are reported as the median of the sampled parameter distributions, with 95% uncertainty intervals (2.5^{th} and 97.5^{th} percentiles) given in parentheses. $\boldsymbol{\varepsilon_{max}}$ (g C MJ^{-1}) denotes the maximum LUE, which is partitioned into sunlit ($\boldsymbol{\varepsilon_{msu}}$) and shaded ($\boldsymbol{\varepsilon_{msh}}$) components in the two-leaf model. VPD_0 (kPa) denotes the sensitivity parameter of the VPD constraint function, whereas VPD_{max} and VPD_{min} (kPa) denote the upper and lower thresholds of the linear VPD scalar used in the LD-MODIS-LUE formulation. For LD-EC-LUE and LD-TL-EC-LUE formulations, T_{min}, T_{max}, and T_{opt} (℃) refer to minimum, maximum, and optimum temperatures for photosynthesis. For LD-MODIS-LUE, T_{min} and T_{max} denote the lower and upper daily-minimum-temperature thresholds used in the LD-MODIS-LUE temperature scalar, where the scalar is constrained to 0 at or below T_{min}, and to 1 at or above T_{max}. Model evaluation includes leave-one-site-out (LOSO) cross-validation and joint calibration across all four EC sites (Global). Model performance is characterized by using the correlation coefficient (*r*) and Root Mean Square Error (RMSE; g C m^{-2} d^{-1}).

Model framework	Literature based parameters	Median with 95% uncertainty intervals	LOSO		Global	
			r	RMSE	*r*	RMSE
LD-EC-LUE	T_{opt} = 28 T_{min} = 2 T_{max} = 48	ε_{max} = 2.43 (1.93, 3.42) VPD_0 = 0.86 (0.44, 1.44)	0.59	1.00	0.61	0.79
LD-TL-EC-LUE	T_{opt} = 28 T_{min} = 2 T_{max}=48	ε_{msu} = 1.63 (1.51, 1.86) ε_{msh} = 2.86 (2.61, 3.18) VPD_0 = 0.95 (0.81, 1.26)	0.63	0.89	0.69	0.64
LD-MODIS-LUE	T_{min} = -8 T_{max} = 9.09	ε_{max} = 1.59 (1.38, 1.86) VPD_{max} = 2.56 (1.82, 4.07) VPD_{min} = 0.37 (0.21, 0.66)	0.63	1.11	0.70	0.60

References:

28. Wang, J. *et al.* An ecologically-constrained deep learning model for tropical leaf phenology monitoring using PlanetScope satellites. *Remote Sens. Environ.* **286**, 113429 (2023).
47. Liu, M., Jiang, P., Chase, J. M. & Liu, X. Global insect herbivory and its response to climate change. *Curr. Biol.* **34**, 2558-2569.e3 (2024).
71. Janssen, T. *et al.* Drought effects on leaf fall, leaf flushing and stem growth in the Amazon forest: reconciling remote sensing data and field observations. *Biogeosciences* **18**, 4445–4472 (2021).
72. Zhang, H., Yuan, W., Dong, W. & Liu, S. Seasonal patterns of litterfall in forest ecosystem worldwide. *Ecol. Complex.* **20**, 240–247 (2014).
74. Maitner, B. S. *et al.* The BIEN R package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods Ecol. Evol.* **9**, 373–379 (2018).
75. Sabatini, F. M. *et al.* sPlotOpen – An environmentally balanced, open-access, global dataset of vegetation plots. *Glob. Ecol. Biogeogr.* **30**, 1740–1764 (2021).
84. Poggio, L. et al. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL* **7**, 217–240 (2021).
85. Zuquim, G. *et al.* Making the most of scarce data: Mapping soil gradients in data-poor areas using species occurrence records. *Methods Ecol. Evol.* **10**, 788–801 (2019).
87. Wan, Z., Hook, S. & Hulley, G. MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V061. (2021) doi:10.5067/MODIS/MOD11A2.061.
93. Muñoz Sabater, J. ERA5-Land monthly averaged data from 1981 to present. *Copernic. Clim. Change Serv. C3S Clim. Data Store CDS* **10**, (2019).
97. Restrepo-Coupe, N. *et al.* LBA-ECO CD-32 Flux Tower Network Data Compilation, Brazilian Amazon: 1999-2006, V2. ORNL Distributed Active Archive Center https://doi.org/10.3334/ORNLDAAC/1842 (2021).
98. Zhang, K. & Zhu, G. A global dataset of terrestrial evapotranspiration (multi-component) and soil moisture (multi-layer) from 1982 to 2022. National Tibetan Plateau Data Center https://doi.org/10.11888/Terre.tpdc.300751 (2025).