# SUPPORTING INFORMATION

## Figure S1

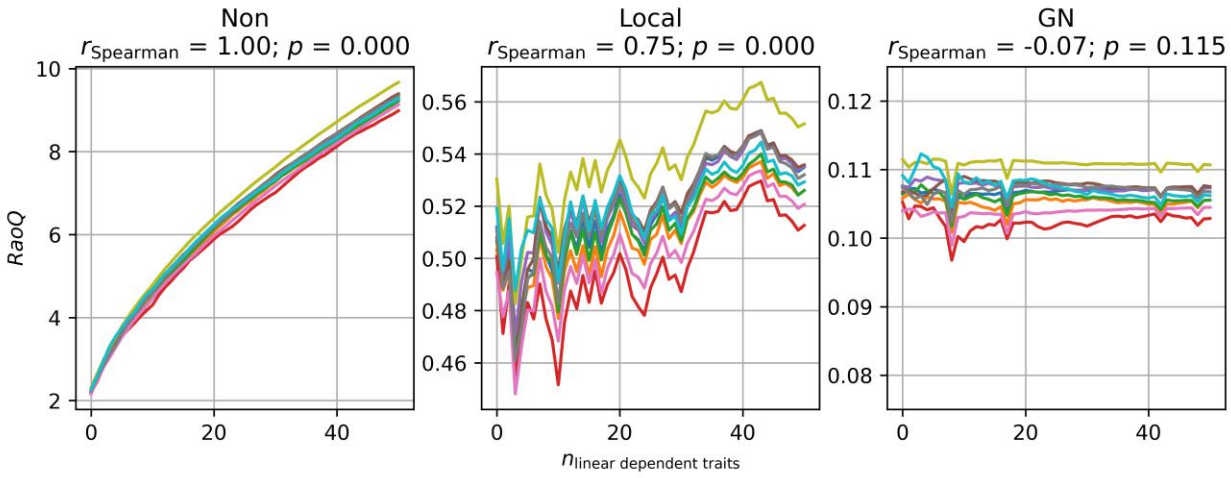


Figure S1. Data-structure independence of generalizable normalization (GN). Twenty synthetic species are used to generate ten different communities (line colors) with abundances drawn from a uniform random distribution in the range [0, 1), and truncated to 0.0 for the cases below 0.5. Frequencies are normalized per community. Three linearly independent traits ($p$-value ≥ 0.30) are generated for the species sampling from a Normal distribution. From these three, up to 50 additional traits are generated by linear combination. Then Gaussian noise of mean 0 and standard deviation 0.01 is added to the traits to represent natural variability and observation uncertainty. Rao quadratic index ($Q$) is calculated for each community using an increasing number of linearly dependent traits. Non and local normalization Rao $Q$ clearly depends on the dataset dimensionality and increases with the number of traits even if these carry no additional information (beyond the one brought by noise). Only GN provides data-structure independent values.

**Supporting S1. Simulation of plant traits.**

Among the vegetation traits of SCOPE model, foliar parameters are the most prone to be correlated (Wright *et al.* 2004; Osnas *et al.* 2013). Therefore, we sampled $N$, $C_{ab}$, $C_{ca}$, $C_{dm}$, and $C_w$ from a conditional Gaussian Mixture Model (GMM) fitted with two spectral libraries used for the development of commonly used leaf radiative transfer models (LOPEX (Hosgood *et al.* 1994) and ANGERS (Feret *et al.* 2008)). These datasets provided estimates of the leaf parameter $N$, an abstraction of the leaf structure correlated with leaf dry matter content ($C_{dm}$) (Jacquemoud & Baret 1990). Since $N$ is not measurable, it is not available in typical leaf trait datasets. The GMM was optimized using the expectation-maximization algorithm (Dempster, Laird & Rubin 1977) implemented in the Python package *Scikit-learn* version 0.22.1 (Pedregosa *et al.* 2011). In addition, we considered that high contents of senescent pigments ($C_s$) and chlorophyll ($C_{ab}$) would be unlikely since the first ones originate from the degradation of photosynthetic pigments and other leaf constituents (Pourcel *et al.* 2007; Mattila *et al.* 2018). High $C_{ab}$ and anthocyanins ($C_{ant}$) contents were considered more likely but not necessarily common cases (Gould 2004; Manetas 2006; Hughes, Morley & Smith 2007). Therefore we used an exponentially decreasing function to scale the randomly sampled values of $C_{ant}$ as a function of $C_{ab}$ (Eq. S1):

$$f_{C,\max} = \mathrm{e}^{z\cdot\left(\frac{100-C_{ab}}{100}-1\right)} \quad \text{(S1.1)}$$

where $z$ was 40 and 7 for senescent pigments and anthocyanins, respectively, and $f_{C,max}$ ranges between 0 and 1.

The remaining structural (Table 1) and soil (Table S3) parameters were simulated without prior knowledge of their covariance or distribution. Therefore, they were randomly sampled within the ranges predefined for "similar" and "dissimilar" species in each pool (see Section 2.1 of the manuscript).

In the simulations, the values of foliar, structural, and soil variables were limited within ranges commonly found in the radiative transfer model literature, either reporting databases (Hosgood *et al.* 1994), model development and parametrization (Feret *et al.* 2008; Féret *et al.* 2017; Vilfan *et al.* 2018), or model inversion (Zhang *et al.* 2016; Bayat, van der Tol & Verhoef 2018; Verhoef, van der Tol & Middleton 2018; Yang *et al.* 2020).

**Table S1**

**Table S1** Correspondence between remote sensing and field biodiversity analysis units in the simulations. The table defines how species, communities, and regions synthesized in species-traits and abundance-species matrices often used in ecological analysis find their corresponding entities in the context of remote sensing studies.

| Field / Ecology | Remote Sensing |
|---|---|
| Species<br><br>Synthetic species are represented as unique sets of plant traits (i.e., vegetation radiative transfer model parameters). In the simulations, we assume no intraspecific variability. Thus, each species is, in practice, a vector of traits allocated into a species-traits matrix row. | Pixel<br><br>In the simulations, we assume imagery able to identify the spectral signature characteristic of each synthetic species (i.e., the reflectance factors predicted for the specie's traits). Given the spatial abstraction, it is not relevant whether this information comes from single or several pixels. |
| Community<br><br>Communities correspond to assemblages of several species sampled from the regional species pool, each featuring a given relative abundance normalized by the sum of the abundances within the community. | Pixels patch or window<br><br>In a remote sensing image, a community would be made of several pixels close in space that could be gathered with different criteria (geographical entities, moving windows, statistical clusters, …). Within this patch, pixels of similar reflectance factors would be associated with individual species (e.g., the spectral species concept from Féret and Asner, (2014)), allowing to determine the species identity and their relative abundance. On occasions, moving windows (e.g., *n*-by-*n* pixels) can be used, assuming each pixel is a species of equivalent relative abundance. |
| Region<br><br>A region would be defined by a set of communities whose biodiversity is separately analyzed (e.g., different rows of the abundance-species matrix), which enables their comparison. Moreover, the joint analysis of these communities allows for quantifying their turnover and the total biodiversity, and therefore the partitioning or analysis of biodiversity across scales. | Image, part, or set of images<br><br>Analogously, several windows or patches of pixels can be analyzed separately and jointly. These can proceed from the same image covering all or part of it, or even from more than one image, depending on the aims and extent of the study and the spatial coverage of the images available. For example, an image could be split into large windows containing smaller moving windows (the former row in this table) analogous to communities. |

**Supporting S2. Radiative transfer modeling in the context of vegetation biodiveristy.**

Radiative transfer models (RTM) represent the interaction of light and matter, predicting the spectral properties (i.e., reflectance factors of the ratio of reflected radiation in the direction of the observer to the illuminating radiation) of leaves and canopies as a function of their biophysical properties. Usually, leaf and canopy are coupled to simulate canopy scale observations. Foliar variables predict leaf reflectance and transmittance factors (Figure S2.1a), which become an input of the canopy model (Figure S2.1b). The second represents the spatial distribution of leaves above the soil and, sometimes, other elements, such as trunks or branches. The simulations represent multiple, contiguous, and narrow spectral bands (hyperspectral) (Figure S2.1c) which can be convolved to coarser spectral bands of satellite sensors (Figure S2.1d). For a deeper view of the radiative transfer modeling of leaves and homogeneous canopies, see Jacquemoud et al. (1990), Verhoef (1984), and Jacquemoud et al. (2009).

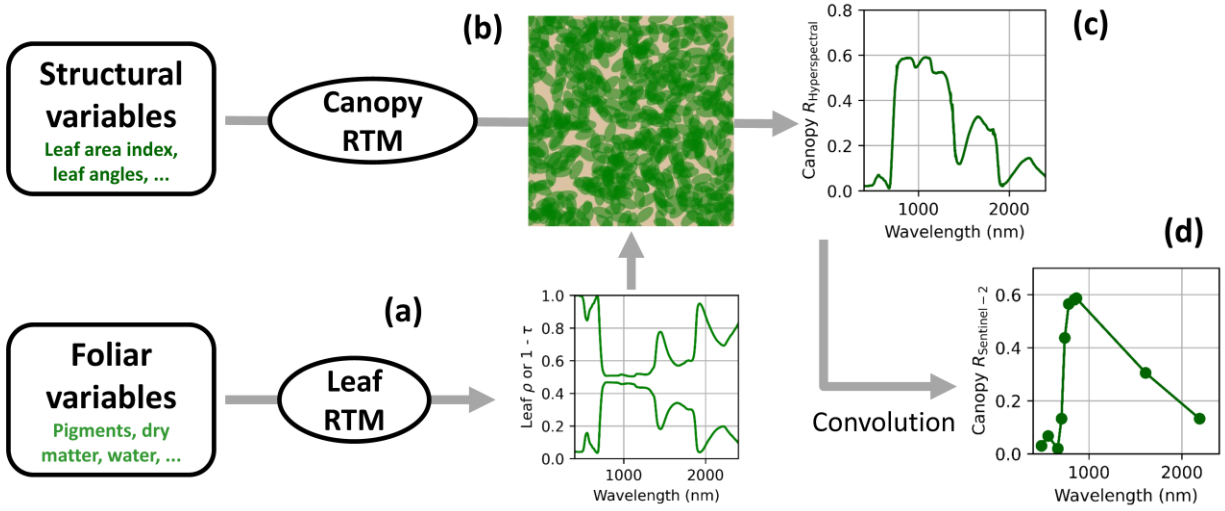


**Figure S2.1.** Schematic representation of a radiative transfer model (RTM) simulation. The leaf RTM predicts foliar reflectance ($\rho$) and transmittance ($\tau$) factors (a). These are inputs of the canopy RTM (b), which predicts the canopy reflectance factor ($R$) for multiple, narrow, and contiguous bands (c). Hyperspectral reflectance factors can then be convolved to the spectral features of satellite sensors (d).

Vegetation foliar and structural variables (inputs of the coupled RTM) can be used as surrogates of plant functional traits to compute functional diversity metrics, which usually rely on the dissimilarity between species traits. The variability of the canopy reflectance factors simulated should accompany the species-related variability of these vegetation variables. Spectral variables can also be used as trait surrogates, and the corresponding diversity can be calculated from the same functional diversity formulations. This approach allows comparing the relationships between the diversities derived from the spectral vegetation and plant traits originating from their variability (Fig. S2.2). For a deeper view of these topics, we recommend the works of Wang et al. (2018), Ma et al. (2020), and Pacheco-Labrador et al.(2022).

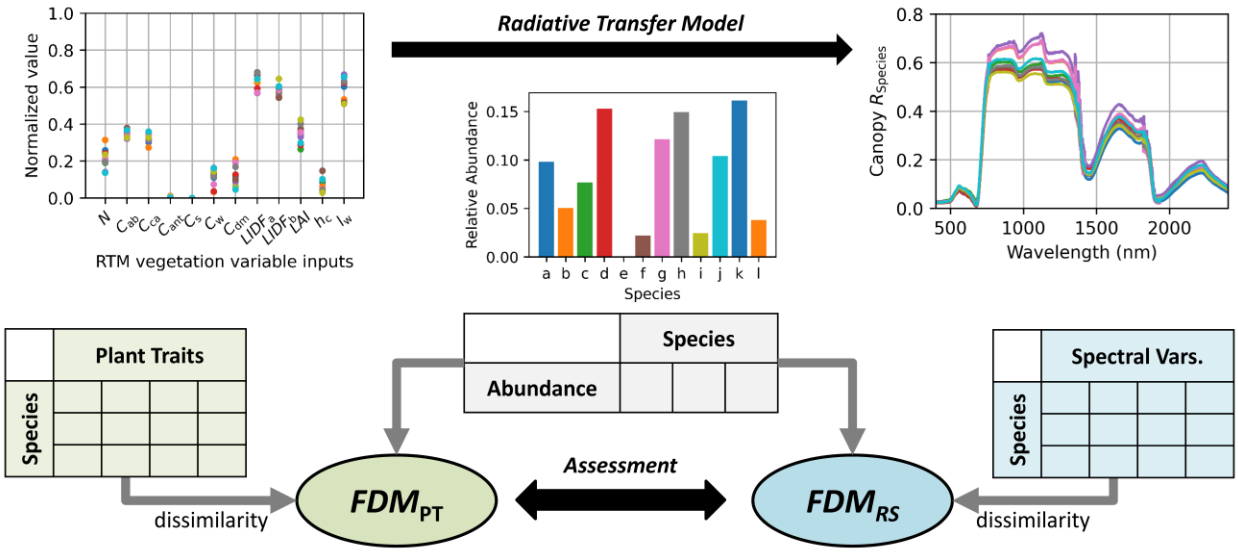


**Figure S2.2.** Summary of the use of radiative transfer models (RTM) to assess the relationships between plant functional and spectral diversity. The RTM vegetation variables used as plant functional trait surrogates are: leaf structural parameter ($N$), chlorophyll ($C_{ab}$), carotenoid ($C_{ca}$), anthocyanins ($C_{ant}$), and senescent pigment ($C_s$) Contents (mass per area), dry matter ($C_{dm}$) and water ($C_w$) contents, leaf angle distribution function parameters ($LIDF_a$ and $LIDF_b$), leaf area index ($LAI$), canopy height ($h_c$), and leaf width ($l_w$).

**Table S2**

**Table S2.** Training and validation statistics of the SCOPE emulator. The table shows dataset sizes, (*n*) presents the mean values of the Root Mean Squared Error (*RMSE*, in reflectance factor units), the relative *RMSE* (*RRMSE*, %), and the normalized *RMSE* (*NRMSE*) of all the bands (400-2400 nm with 1 nm step).

| Dataset | *n (*samples*)* | *RMSE* (-) | RRMSE (%) | NRMSE (%) |
|---|---|---|---|---|
| Training | 6000 | 0.0050 | 5.92 | 0.95 |
| Validation | 1000 | 0.0056 | 6.59 | 1.29 |

**Table S3**

**Table S3.** Additional SCOPE model parameters, symbols, units, and bounds commonly found in the literature (see supporting 1) used in the simulation.

| Parameter | Symbol | Units | Bounds |
|---|---|---|---|
| Soil Parameters | | | |
| Soil brightness | $B$ | - | [0.5, 1.0] |
| Spectral shape "latitude" | $Lat$ | deg | [20, 40] |
| Spectral shape "longitude" | $Lon$ | deg | [45, 65] |
| Soil moisture capacity | $SMC$ | % | [5, 55] |
| Soil moisture content | $SM_p$ | - | [0, 1] |
| Sun view and atmosphere | | | |
| Sun zenith angle | $\theta_{sun}$ | deg | 30 |
| Diffuse-to-global radiation ratio | $\delta_{DG}$ | - | 0.2 |

65 **Figure S2**

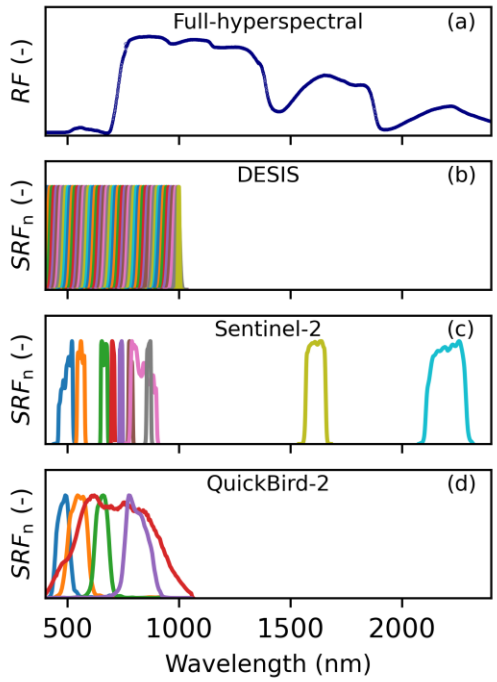


Figure S2. Select of spectral configurations representing the main types of optical remote sensing missions available. The selection covers the combinations of hyperspectral-multispectral and VNIR-VSWIR types of imagers. The figure presents the full-hyperspectral configuration simulating a canopy reflectance factor (*RF*) (a) and the normalized Spectral Response Function ($SRF_n$) of the remote sensing imagers DESIS (b), Sentinel-2 (c), and QuickBird-2 (d).

66

67 **Figure S3**

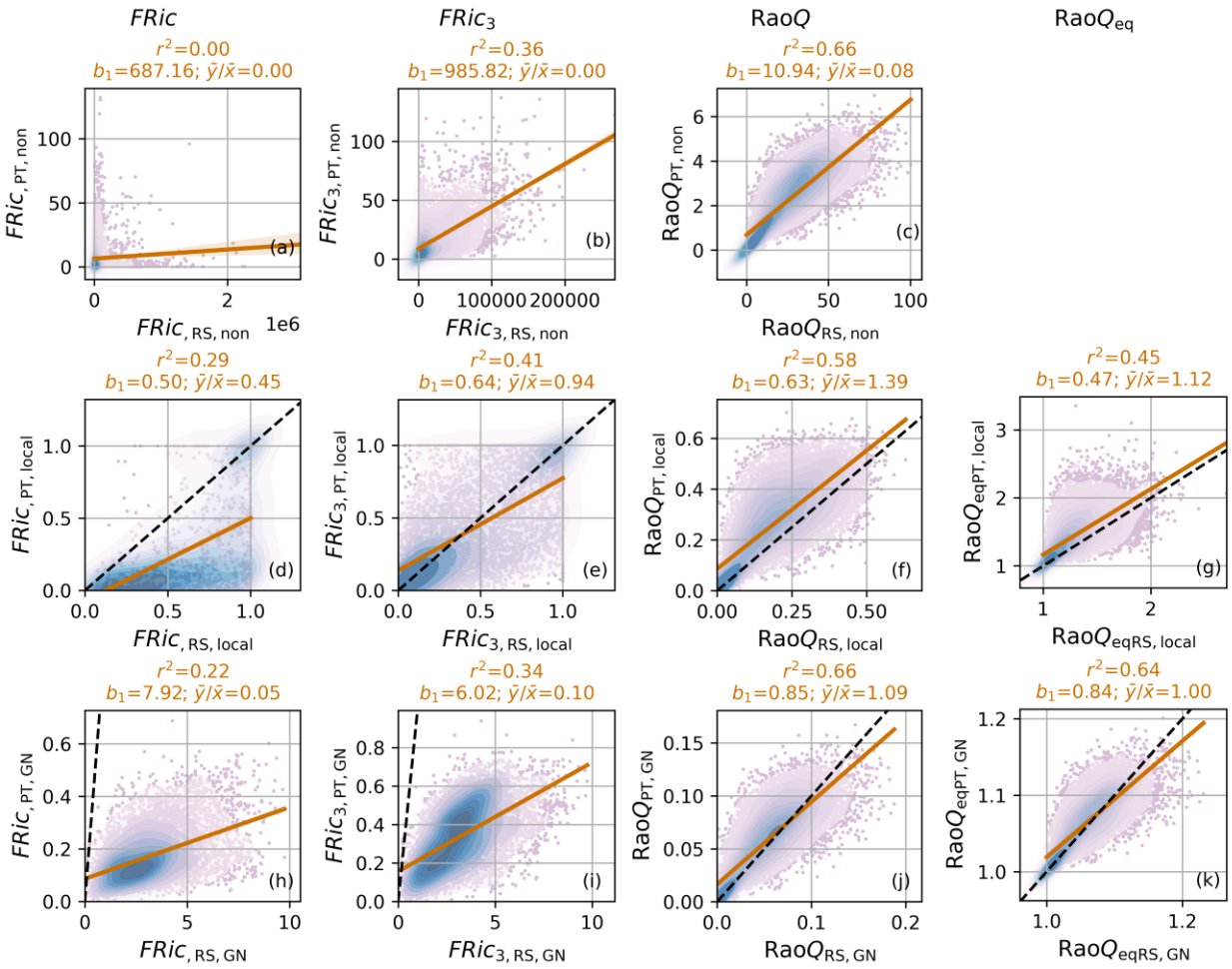


Figure S3. Analogous to Fig. 3, simulating a maximum of 30 species. Diversity indices and equivalent numbers computed with non (first row), local (second row), and generalizable normalization (GN, third row) corresponding to field plant trait ($y$-axis) and full-hyperspectral remote sensing ($x$-axis) datasets. *FRic* (first column), *FRic* limited to up to 3 principal components (second column), Rao $Q$ index (third column) and its equivalent number (fourth column). Squared Pearson correlation coefficient ($r^2$), the linear model slope ($b_1$), and the averages' ratio ($\bar{y}/\bar{x}$) are presented.

68 **Figure S4**

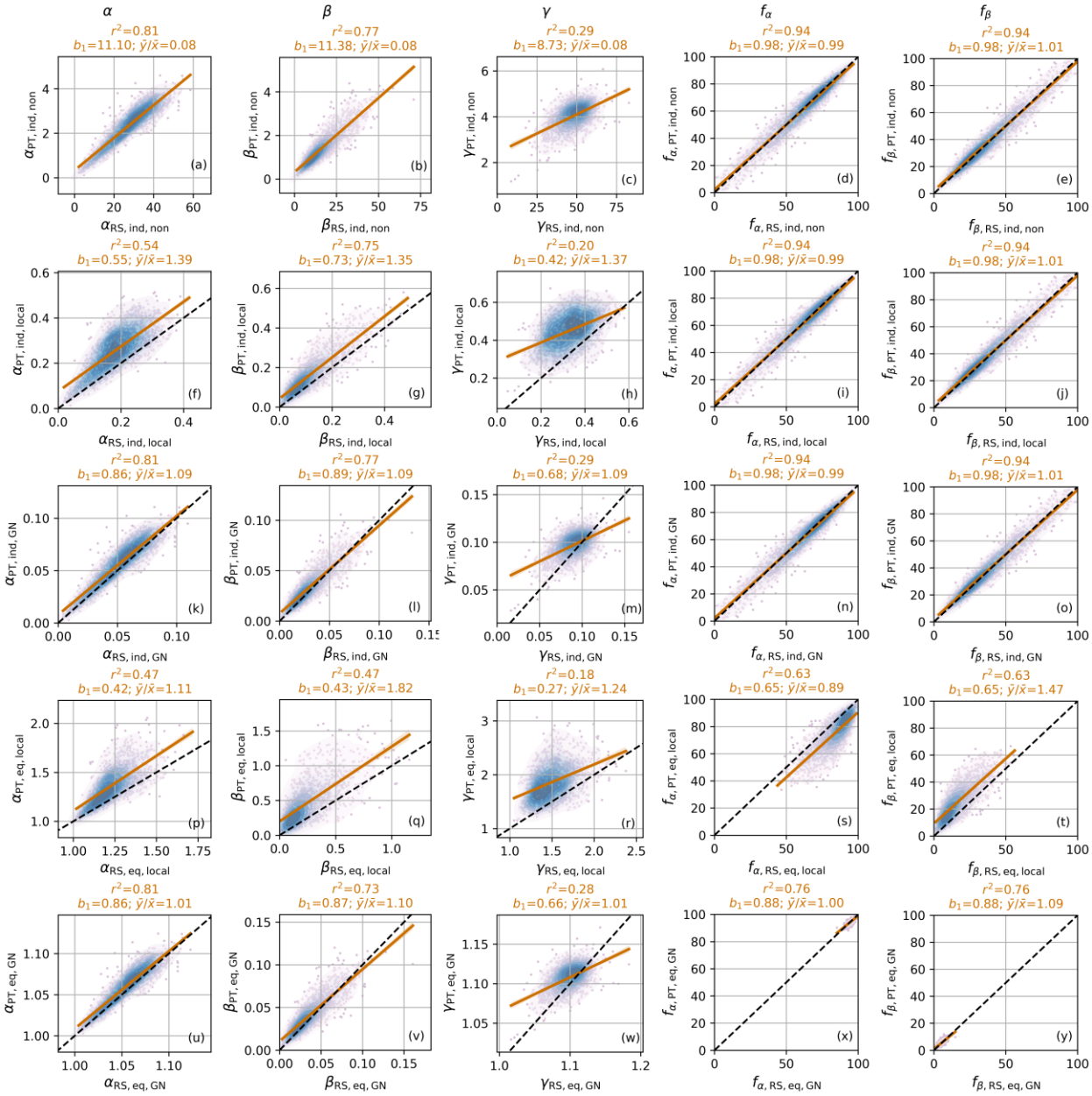


Figure S4. Analogous to Fig. 4, simulating a maximum of 30 species. Diversity decomposition using the diversity index Rao $Q$ with non (first row), local (second row), and generalizable normalization (GN, third row), or its equivalent number with local (fourth row), and GN (fifth row) corresponding to field plant trait ($y$-axis) and full-hyperspectral remote sensing ($x$-axis) datasets. Estimates of $\alpha$ (first column), $\beta$ (second column), and $\gamma$ (third column) diversity, as well as the fractions of $\alpha$ (fourth column) and $\beta$ (fifth column). Squared Pearson correlation coefficient ($r^2$), the linear model slope ($b_1$), and the averages' ratio ($\overline{y}/\overline{x}$) are presented.

69

**Figure S5**

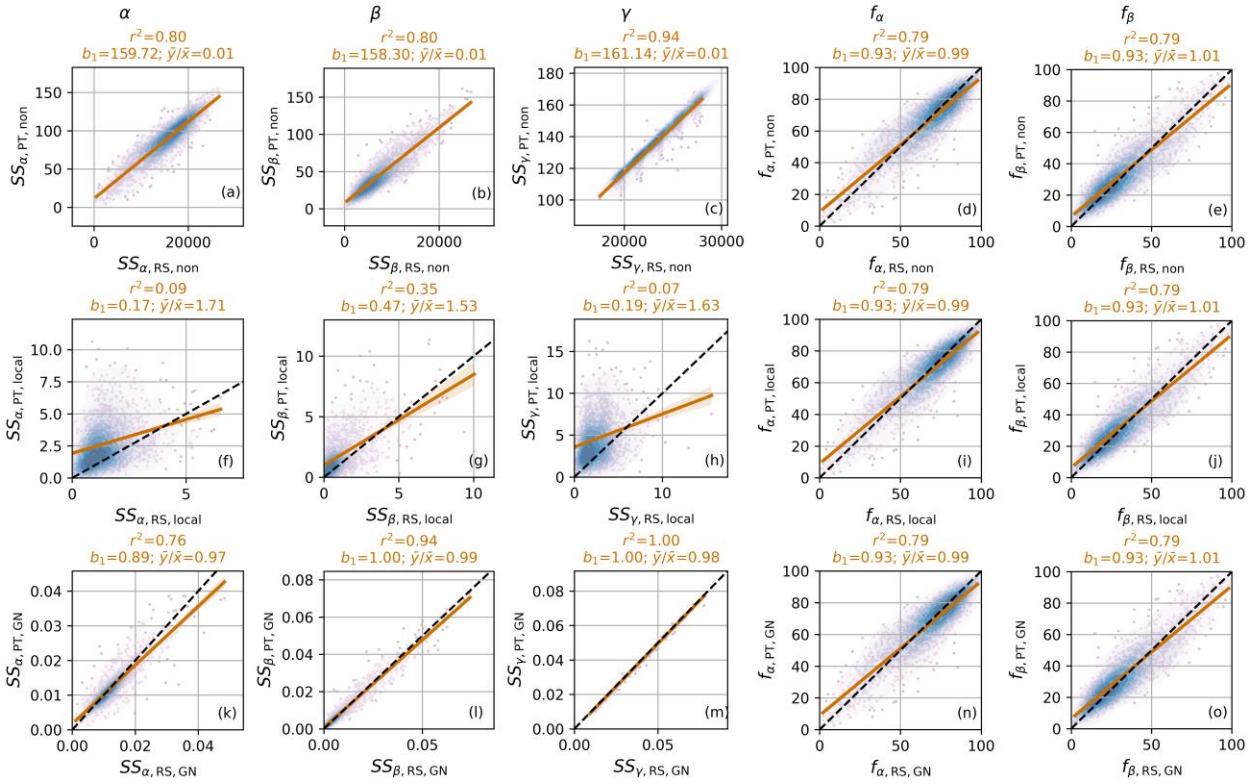


Figure S5. Analogous to Fig. 5, simulating a maximum of 30 species. Diversity decomposition using variance without (first row), with local (second row), and generalizable normalization (GN, third row) corresponding to field plant trait ($y$-axis) and full-hyperspectral remote sensing ($x$-axis) datasets. Estimates of $\alpha$ (first column), $\beta$ (second column), and $\gamma$ (third column) components, as well as the fractions of $\alpha$ (fourth column) and $\beta$ (fifth column). Squared Pearson correlation coefficient ($r^2$), the linear model slope ($b_1$), and the averages' ratio ($\bar{y}/\bar{x}$) are presented.

## REFERENCES


Bayat, B., van der Tol, C. & Verhoef, W. (2018) Integrating satellite optical and thermal infrared observations for improving daily ecosystem functioning estimations during a drought episode. *Remote Sensing of Environment,* **209,** 375-394.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological),* **39,** 1-38.

Féret, J.-B. & Asner, G.P. (2014) Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecological Applications,* **24,** 1289-1296.

Feret, J.-B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidel, L.P.R., Ustin, S.L., le Maire, G. & Jacquemoud, S. (2008) PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sensing of Environment,* **112,** 3030-3043.

Féret, J.B., Gitelson, A.A., Noble, S.D. & Jacquemoud, S. (2017) PROSPECT-D: Towards modeling leaf optical properties through a complete lifecycle. *Remote Sensing of Environment,* **193,** 204-215.

Gould, K.S. (2004) Nature's Swiss Army Knife: The Diverse Protective Roles of Anthocyanins in Leaves. *Journal of biomedicine & biotechnology,* **2004,** 314-320.

Hosgood, B., Jacquemoud, S., Andreoli, G., Verdebout, J., Pedrini, G. & Schmuck, G. (1994) Leaf Optical Properties EXperiment 93 (LOPEX93). (ed. E.C.J.R. Centre). Ispra (Italy). Available online: [http://opticleaf.ipgp.fr/databases/LDB_lopex1993.xls], last accessed 11-Nov-2020.

Hughes, N.M., Morley, C.B. & Smith, W.K. (2007) Coordination of anthocyanin decline and photosynthetic maturation in juvenile leaves of three deciduous tree species. *New Phytologist,* **175,** 675-685.

Jacquemoud, S. & Baret, F. (1990) PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment,* **34,** 75-91.

Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C. & Ustin, S.L. (2009) PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sensing of Environment,* **113,** S56-S66.

Ma, X., Migliavacca, M., Wirth, C., Bohn, J.F., Huth, A., Richter, R. & Mahecha, D.M. (2020) Monitoring Plant Functional Diversity Using the Reflectance and Echo from Space. *Remote Sensing,* **12**.

Manetas, Y. (2006) Why some leaves are anthocyanic and why most anthocyanic leaves are red? *Flora - Morphology, Distribution, Functional Ecology of Plants,* **201,** 163-177.

Mattila, H., Valev, D., Havurinne, V., Khorobrykh, S., Virtanen, O., Antinluoma, M., Mishra, K.B. & Tyystjärvi, E. (2018) Degradation of chlorophyll and synthesis of flavonols during autumn senescence—the story told by individual leaves. *AoB PLANTS,* **10**.

Osnas, J.L.D., Lichstein, J.W., Reich, P.B. & Pacala, S.W. (2013) Global Leaf Trait Relationships: Mass, Area, and the Leaf Economics Spectrum. *Science,* **340,** 741.

Pacheco-Labrador, J., Migliavacca, M., Ma, X., Mahecha, M.D., Carvalhais, N., Weber, U., Benavides, R., Bouriaud, O., Barnoaiea, I., Coomes, D.A., Bohn, F.J., Kraemer, G., Heiden, U., Huth, A. & Wirth, C. (2022) Challenging the link between functional and spectral diversity with radiative transfer modeling and data. *Remote Sensing of Environment,* **280,** 113170.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, É. (2011) Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research,* **12,** 2825-2830.

Pourcel, L., Routaboul, J.-M., Cheynier, V., Lepiniec, L. & Debeaujon, I. (2007) Flavonoid oxidation in plants: from biochemical properties to physiological functions. *Trends in Plant Science,* **12,** 29-36.

Verhoef, W. (1984) Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment,* **16,** 125-141.

Verhoef, W., van der Tol, C. & Middleton, E.M. (2018) Hyperspectral radiative transfer modeling to explore the combined retrieval of biophysical parameters and canopy fluorescence from FLEX – Sentinel-3 tandem mission multi-sensor data. *Remote Sensing of Environment,* **204,** 942-963.

Vilfan, N., Van der Tol, C., Yang, P., Wyber, R., Malenovský, Z., Robinson, S.A. & Verhoef, W. (2018) Extending Fluspect to simulate xanthophyll driven leaf reflectance dynamics. *Remote Sensing of Environment,* **211,** 345-356.

Wang, R., Gamon, J.A., Cavender-Bares, J., Townsend, P.A. & Zygielbaum, A.I. (2018) The spatial sensitivity of the spectral diversity–biodiversity relationship: an experimental test in a prairie grassland. *Ecological Applications,* **28,** 541-556.

Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, Ü., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. & Villar, R. (2004) The worldwide leaf economics spectrum. *Nature,* **428,** 821.

Yang, P., van der Tol, C., Yin, T. & Verhoef, W. (2020) The SPART model: A soil-plant-atmosphere radiative transfer model for satellite measurements in the solar spectrum. *Remote Sensing of Environment,* **247,** 111870.

Zhang, Y., Guanter, L., Berry, J.A., van der Tol, C., Yang, X., Tang, J. & Zhang, F. (2016) Model-based analysis of the relationship between sun-induced chlorophyll fluorescence and gross primary production for remote sensing applications. *Remote Sensing of Environment,* **187,** 145-155.