Satellite remote sensing of plant functional diversity

Hauser, L.T.

Chapter 7.: Supplementary Materials

Fig. S. 1. Field photography illustrative of different locations and land use types

Forested land use (Oak forest: left, Mixed pine forest: right)

Shrubland (Mixed with Holm oak: left, Cistus spp. shrubs: right)

Chestnut Plantations (managed: left, abandoned: right)

S. 2: Cross-validation of Plant Area Index

We cross-validated the Plant Area Index (PAI) measurements from hemispherical photography with above and below canopy measurements taken with Photosynthetically Active Radiation (PAR) sensor (Apogee MQ-301; handheld device), quantifying the relative quantity of incident solar radiation absorbed by vegetation. PAR measurements were taken above and below the canopy to calculate the canopy absorbed light. A strong correlation was been found between PAR measurements and the PAI derived hemispherical photography. PAR measurements, however, did show high susceptibility to cloud cover and changing weather conditions resulting in a number of outliers as also indicated by Garrigues et al. (2008).

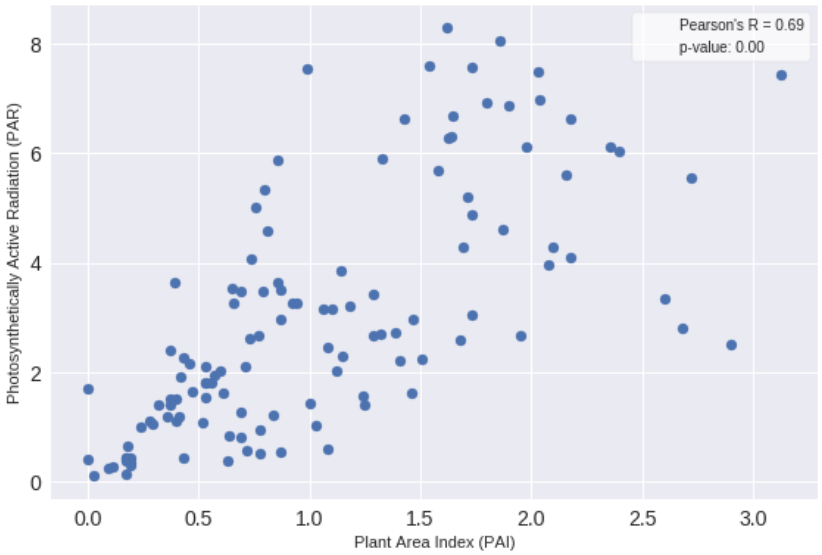


Fig. S. 2a: Comparison of Plant Area Index (PAI) measurements derived from hemispherical photography against Photosynthetically Active Radiation (PAR) measurements.

Further cross-validation was performed through comparison of PAI measurements against the fraction of vegetation/soil through spectral unmixing. Spectral unmixing was conducted using the HSDAR package in R (Lehnert et al., 2018). The principles behind the linear spectral unmixing approach are based on Sohn and McCoy (1997). Two endmember signals were considered in the unmixing exercise; the in-situ soil spectral reflectance, and the leaf spectral reflectance signals generated by running the PROSPECT RTM in forward mode based on in-situ measured field traits. To facilitate the former, we collected handheld hyperspectral reflectance data of representative soils in the study area using the RS-3500 spectroradiometer (350-2500 nm, 8 nm resolution) developed by Spectral Evolution.

The spectral unmixing analysis resulted in a fraction (%) of leaf vegetation signal, a fraction (%) of the soil spectral signal and an 'error' signal (%) that may indicate that endmember spectra do not fit well to the predefined endmember spectra. The fraction of vegetation was used to cross-validate the Plant Area Index measurements as an indicator of the vegetation density. As expected, high PAI strongly correlates with a high fraction of vegetation, inversely this results in a negative correlation with the fraction of soil in the spectral signal.

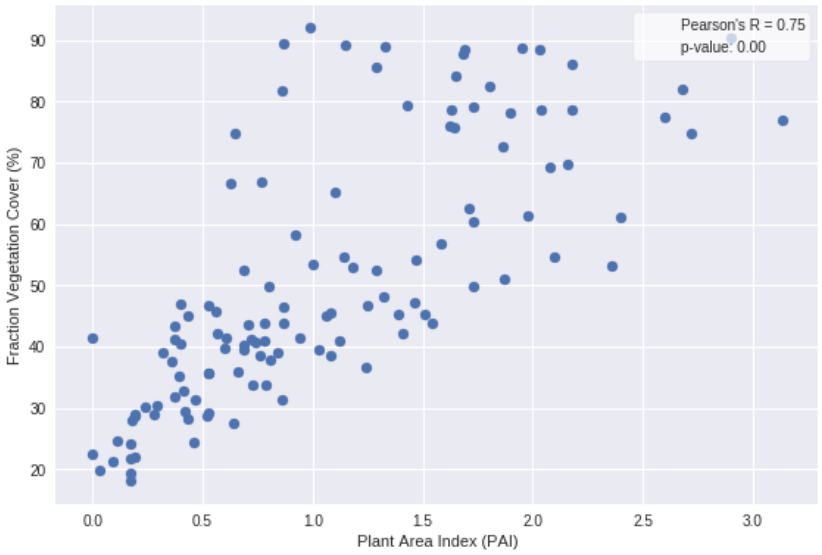


Fig. S. 2b: Comparison of Plant Area Index (PAI) measurements derived from hemispherical photography against Fraction of Vegetation Cover (%) derived from spectral unmixing

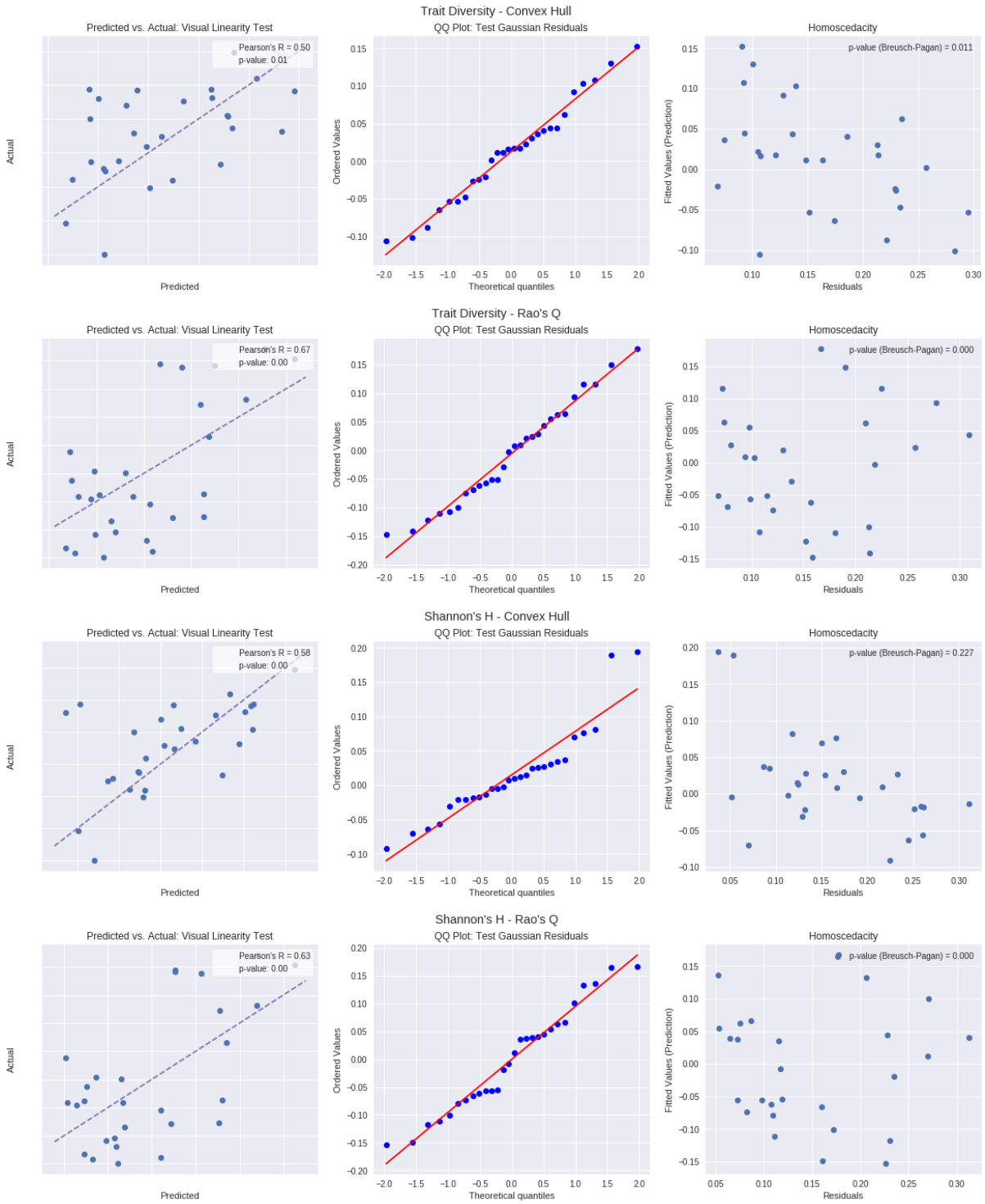
Table S. 3: Spectral characteristics of the Sentinel-2 bands used

Acronym	Central (nm)	Width (nm)	Spatial resolution (m)
B2	490	65	10
B3	560	35	10
B4	665	30	10
B5	705	15	20
B6	740	15	20
B7	783	20	20
B8a	865	20	20
B11	1610	90	20
B12	2190	180	20

Table S. 4: Overview of the configuration of input parameters using in PROSAIL-simulated spectral diversity which includes the two SAIL implementations; with fixed parameters and with variable (Genetic algorithm-optimized) parameter values.

Model	Parameter	Unit	Implementation: Fixed SAIL parameters	Implementation: Variable SAIL parameters
PROSPECT	Leaf structure	-	Fixed: 1.6	Fixed: 1.6
	Chlorophyll a+b	ug/cm^2	In-situ measurements	In-situ measurements
	Carotenoids	ug/cm^2	In-situ measurements	In-situ measurements
	Equivalent water thickness	g/cm^2	In-situ measurements	In-situ measurements
	Dry matter content	g/cm^2	In-situ measurements	In-situ measurements
	Brown pigment	-	Fixed: 0.01	Fixed: 0.01
SAIL	Leaf area index	m2 / m2	In-situ measurements: PAI-derived	In-situ measurements: PAI-derived
	Average leaf angle	o	Fixed: Planophile	Search range: 20-85
	Hotspot	-	Fixed: 0.01	Fixed: 0.01
	Dry/wet soil ratio	-	Fixed study area soil reflectance	Search range: 0-1
	Soil brightness	-	Fixed study area soil reflectance	Search range: 0-2
Positional	Solar zenith	o	Sentinel-2 derived	Sentinel-2 derived
	Observer zenith	o	Sentinel-2 derived	Sentinel-2 derived
	Relative azimuth	o	Sentinel-2 derived	Sentinel-2 derived

Fig. S. 5: Review of assumptions for linear mixed-effect regression analyses: Linearity, Normality of residuals, Homoscedasticity



S. 6: Linear mixed-effect models; the role of plant diversity, vegetation cover, and landscape morphology in explaining Sentinel-2 spectral diversity observations

Table S. 6a: Linear mixed-effect model summary and relative weights analysis with spectral diversity calculated based on the Convex Hull Volume (CHV) as the dependent variable and vegetation cover, landscape morphology, and CHV-based calculation of trait diversity as predictors, while location is set as the random effect. τ00 - random factor intercept variance (i.e. between-subject variance of the location), ICC – random factor intraclass correlation coefficient

	Spectral Diversity (CHV)					
Predictors	*Estimates*	*CI*	*β*	*p*	*VIF*	*Relative weights*
(Intercept)	0.10	0.01 – 0.19	**0.00**	**0.035**	-	-
Trait diversity (CHV)	0.16	-0.19 – 0.51	0.19	0.332	2.60	6.36
Vegetation Cover	0.95	0.19 – 1.72	**0.59**	**0.019**	1.29	84.46
Landscape Morphology	0.72	-0.60 – 2.03	0.24	0.254	2.19	9.18
Random Effects						
σ^2	0.00					
$\tau_{00\ Location}$	0.00					
ICC	0.00					
$N_{Location}$	14					
Observations	28					
$R^2_{Marginal}$ / $R^2_{Conditional}$	0.221 / 0.221					

Table S. 6b: Linear mixed-effect model and relative weights analysis summary with spectral diversity calculated based on the Rao's quadratic entropy (Rao's Q) as the dependent variable and vegetation cover, landscape morphology, and Rao's Q-based trait diversity as predictors, while location is set as the random effect. τ00 - random factor intercept variance (i.e. between-subject variance of the location), ICC – random factor intraclass correlation coefficient

	Spectral Diversity (Rao's Q)					
Predictors	*Estimates*	*CI*	*β*	*p*	*VIF*	*Relative weights*
(Intercept)	-0.04	-0.16 – 0.08	0.00	0.508	-	-
Trait diversity (Rao's Q)	0.25	-0.12 – 0.62	0.24	0.161	1.71	5.83
Vegetation Cover	2.07	0.86 – 3.29	**0.73**	**0.003**	1.20	63.23
Landscape Morphology	2.90	0.72 – 5.08	**0.56**	**0.014**	1.52	30.93
Random Effects						
σ^2	0.01					
$\tau_{00\ Location}$	0.00					
ICC	0.18					
$N_{Location}$	14					
Observations	28					
$R^2_{Marginal}$ / $R^2_{Conditional}$	0.400 / 0.507					

Table S. 6c: Linear mixed-effect model summary and relative weights analysis with spectral diversity calculated based on the Convex Hull Volume (CHV) as the dependent variable and vegetation cover, landscape morphology, and Shannon's H taxonomic diversity as fixed effect predictors, while location is set as the random effect. τ00 - random factor intercept variance (i.e. between-subject variance of the location), ICC – random factor intraclass correlation coefficient

	Spectral Diversity (CHV)					
Predictors	*Estimates*	*CI*	*β*	*p*	*VIF*	*Relative weights*
(Intercept)	0.10	0.04 – 0.16	**0.00**	**0.003**	-	-
Shannon's H	0.24	0.02 – 0.45	**0.40**	**0.032**	2.19	42.24
Vegetation Cover	0.86	0.20 – 1.52	**0.53**	**0.015**	1.28	52.98
Landscape Morphology	0.53	-0.67 – 1.73	0.18	0.354	1.84	4.78
Random Effects						
σ^2	0.00					
$\tau_{00\ Location}$	0.00					
ICC	0.00					
$N_{\ Location}$	14					
Observations	28					
$R^2_{\ Marginal}$ / $R^2_{\ Conditional}$	0.342 / 0.342					

Table S. 6d: Linear mixed-effect model summary and relative weights analysis with spectral diversity calculated based on the Rao's quadratic entropy (Rao's Q) as the dependent variable and vegetation cover, landscape morphology, and Shannon's H taxonomic diversity as fixed effect predictors, while location is set as the random effect. τ00 - random factor intercept variance (i.e. between-subject variance of the location), ICC – random factor intraclass correlation coefficient

Spectral Diversity (Rao's Q)						
Predictors	*Estimates*	*CI*	*β*	*p*	*VIF*	*Relative weights*
(Intercept)	0.01	-0.11 – 0.13	0.00	0.916	-	-
Shannon's H	0.11	-0.32 – 0.54	0.16	0.585	2.19	2.98
Vegetation Cover	1.81	0.48 – 3.13	**0.63**	**0.012**	1.28	66.53
Landscape Morphology	2.39	-0.03 – 4.82	0.46	0.053	1.84	30.49
Random Effects						
σ^2	0.01					
$\tau_{00\ Location}$	0.00					
ICC	0.33					
$N_{\ Location}$	14					
Observations	28					
$R^2_{\ Marginal}$ / $R^2_{\ Conditional}$	0.327 / 0.552					

S. 7. Isolating biodiversity signals in spectral diversity

Spectral diversity approaches tend to circumvent the challenges and difficulties in isolating direct biodiversity signals, e.g. functional traits or species discrimination, in the spectral signal. This makes spectral diversity relatively straightforward to deploy. However, as shown here, the approach risks the spectral diversity signal to be dominated by other factors and, in the end, only poorly relate to in-situ plant diversity. Our ability to account for spectrally dominant confounding factors, such as vegetation cover, could improve the robustness and relationship of spectral diversity approaches to in-situ plant diversity across spatial, spectral, geographic, and temporal resolutions. We suggest three main techniques that can be integrated or complement spectral diversity indices to control for components of vegetation cover including the abundance of soil signals, variation in vegetation density, and canopy architecture (morphological traits) in multispectral satellite-based observations. These techniques include spectral unmixing techniques, radiative transfer model inversion, and data fusion (Musavi et al., 2015).

The separation of the soil signal from the effects of vegetation (morphology and biochemistry) can be done through spectral unmixing (Asner and Heidebrecht, 2002; Clasen et al., 2015). Gholizadeh et al. (2018) proposed a metric of spectral diversity that is soil abundance normalized and applicable to high-resolution observations to allow for the removal of 'pure' soil pixels. In coarser satellite observations, pixels are commonly mixed aggregates of both soil and vegetation spectra. In those cases, spectral unmixing and the normalization of the fraction of soil signal can help limit the confounding role of vegetation cover to some degree. Yet, representative and extensive spectral libraries are required to go beyond coarse soil/vegetation fractional patterns to assess the role of different growth forms and canopy architecture (Nidamanuri and Ramiya, 2013).

RTMs, specifically those intended for (heterogeneous) canopies, partition and parameterize the spectral reflectance of canopies into different layers and components of soil background and vegetation (Jacquemoud et al. 2009; Myneni et al. 1995). In figures 2-3, we applied *forward* radiative transfer modeling to simulate leaf and canopy spectra. The *inversion* of such models, on the other hand, enables the retrieval of plant traits from spectral information which are of specific interest for functional diversity assessments. This could allow to assess and adjust spectral diversity indices in consideration of soil signal abundance, and canopy architecture. Alternatively, it can facilitate a direct comparison of spectrally retrieved leaf trait values against in-situ leaf trait measurements (e.g. Ali et al., 2020a, 2020b; Brown et al., 2019; Rossi et al., 2020). In addition, recent advances have proposed the possibility of leaf-canopy models using the Directional Area Scattering Factor to specifically correct spectral data for canopy structural effects (Adams et al., 2018). Although still in their infancy, these models could facilitate targeted decoupling of spectra in canopy structure and biochemistry signals.

The inversion of canopy RTMs to estimate plant traits and biophysical parameters from spectra and radiometry is far from trivial (Musavi et al., 2015). It relies on a parameterization of the canopy structure characteristics (Table S. 4). The inversion is often ill-posed and prone to a range of equally possible solutions which challenges the correct simultaneous estimation of the parameters, especially in the case of multispectral information (Combal et al., 2003; Koetz et al., 2007). The forthcoming launch of hyperspectral satellite imagers (e.g. EnMAP, SBG,

CHIME; Cavender-Bares, Gamon, and Townsend 2020) offers prospective on more detailed spectral information that might benefit inversion exercises.

Alternatively, empirical relationships between different spectra and plant traits can be established through the use of machine learning, Partial Least Squares Regression, and spectral vegetation indices (Durán et al., 2019; Ma et al., 2019; Schneider et al., 2017; Verrelst et al., 2019a, 2015). These methods, however, require large training datasets which hampers their scalability and vegetation indices have been found to be site- and time-specific with limited generality, and potentially relevant spectral information is disregarded (Ali et al., 2020a; Verrelst et al., 2015)

Ancillary data (e.g. on soil and vegetation) and data assimilation can help overcome the ill-posedness in solving RTM inversion by constraining canopy models with a priori information (Lahoz and Schneider, 2014; Lewis et al., 2012; Rivera et al., 2013). In our study, we relied on forward modelling by including in-situ measurements to generate simulated leaf and canopy spectra, but those will not be available for large-scale applications. The ingestion of LIDAR data is interesting to complement pixel-based optical reflectance for better characterization of the canopy structure including its morphological traits and vertical profile (Hakkenberg et al., 2018; Koetz et al., 2007). Such information can help further understand the role of vegetation cover as a confounding factor (Combal et al., 2003).

Fig. S. 8. Field photography illustrative of different locations and land use types

Forested locations	**Shrubland locations**
Oak forest:	Holm oak:
	
Mixed forest:	Mixed shrubland:
	
Mixed pine forest:	Cistus spp. Shrubs:
	

S. 9: Cross-validation of Leaf Area Index (LAI) against Photosynthetically Active Radiation (PAR) measurements

We cross-validated the Leaf Area Index (LAI) measurements from hemispherical photography with above and below canopy measurements taken with Photosynthetically Active Radiation (PAR) sensor (Apogee MQ-301; handheld device), quantifying the relative quantity of incident solar radiation absorbed by vegetation. PAR measurements were taken above and below the canopy to calculate the canopy absorbed light. A strong correlation was been found between PAR measurements and the LAI-derived hemispherical photography. PAR measurements however did show high susceptibility to cloud cover and changing weather conditions resulting in a number of outliers as also indicated by Garrigues et al. (2008).

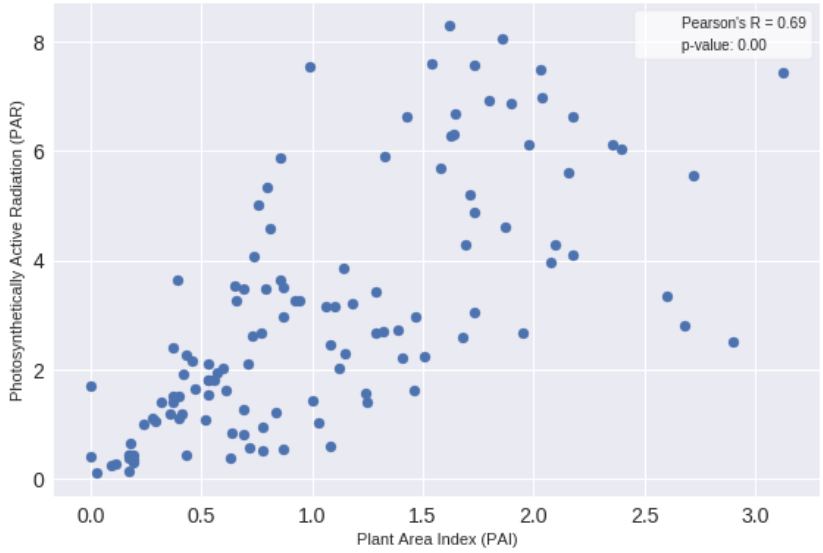


Fig. S. 9: Comparison of LAI Area Index (LAI) measurements derived from hemispherical photography against Photosynthetically Active Radiation (PAR) measurements.

S. 10: Plot-wise linear correlation between satellite-based single trait estimates and in-situ field measurements

Table S. 10: Overview of bivariate correlation (Pearson's r) and coefficient of determination (R2) between plot-wise satellite-based trait estimates and in-situ measured trait observations. Root Mean Square Error (RMSE) quantifies the deviations in estimated traits. Moran's I is indicative of the spatial association in the residuals of neighbouring points. Moran's I null-hypothesis assumes spatial independence.

Pixel-level trait estimates					
Algorithm	Trait	Pearson's R	R^2	RMSE	Moran's I
SVR (Hybrid)	Leaf Area Index (m2/m2)	0.76**	0.58	0.49	0.73**
	Leaf Mass per Area (mg/cm²)	0.81**	0.66	6.36	0.69**
	Leaf Chlorophyll (µg/cm²)	0.56**	0.31	7.07	0.63**
ARTMO (LUT)	Leaf Area Index (m2/m2)	0.79**	0.63	0.51	0.64**
	Leaf Mass per Area (mg/cm²)	0.82**	0.68	31.06	0.69**
	Leaf Chlorophyll (µg/cm²)	0.30**	0.09	12.03	0.68**
SNAP	Leaf Area Index (m2/m2)	0.74**	0.55	0.62	0.73**
	Leaf Mass per Area (mg/cm²)	0.71**	0.51	5.94	0.69**
	Leaf Chlorophyll (µg/cm²)	0.43**	0.19	13.79	0.69**
**: Significant correlation (p <0.01)					

S. 11: Plot-wise regression between satellite-based single trait estimates and in-situ field measurements based on spatial error modelling

Table S. 11. Overview of spatial error model regression between plot-wise satellite-based trait estimates and in-situ measured trait observations (Coefficient, St.Error and Nagelkerke's Pseudo-R2). Root Mean Square Error (RMSE) quantifies the deviations in estimated traits. Moran's I is indiciative of the spatial association of in the residuals neigboring points. Moran's I null-hypothesis assumes spatial independence.

Pixel-level trait estimates					
Algo-rithm	Trait	Spatial Error Model			Moran's I
		Estimate	St.Error	Pseudo-R^2	
SVR (Hybrid)	Leaf Area Index (m2/m2)	0.30**	0.04	0.93	0.43ns
	Leaf Mass per Area (mg/cm²)	0.33**	0.10	0.91	0.41ns
	Leaf Chlorophyll (µg/cm²)	0.17**	0.06	0.76	0.43ns
ARTMO (LUT)	Leaf Area Index (m2/m2)	0.51**	0.07	0.86	0.43ns
	Leaf Mass per Area (mg/cm²)	1.45**	0.28	0.89	0.42ns
	Leaf Chlorophyll (µg/cm²)	0.15ns	0.14	0.35	0.41ns
SNAP	Leaf Area Index (m2/m2)	0.38**	0.05	0.93	0.43ns
	Leaf Mass per Area (mg/cm²)	0.06**	0.04	0.81	0.37ns
	Leaf Chlorophyll (µg/cm²)	0.08**	0.06	0.88	0.41ns

**: Significant correlation (p <0.01), ns: Not significant (p>0.05)

S. 12: SNAP biophysical processor retrieval of single trait estimates compared against field measurements

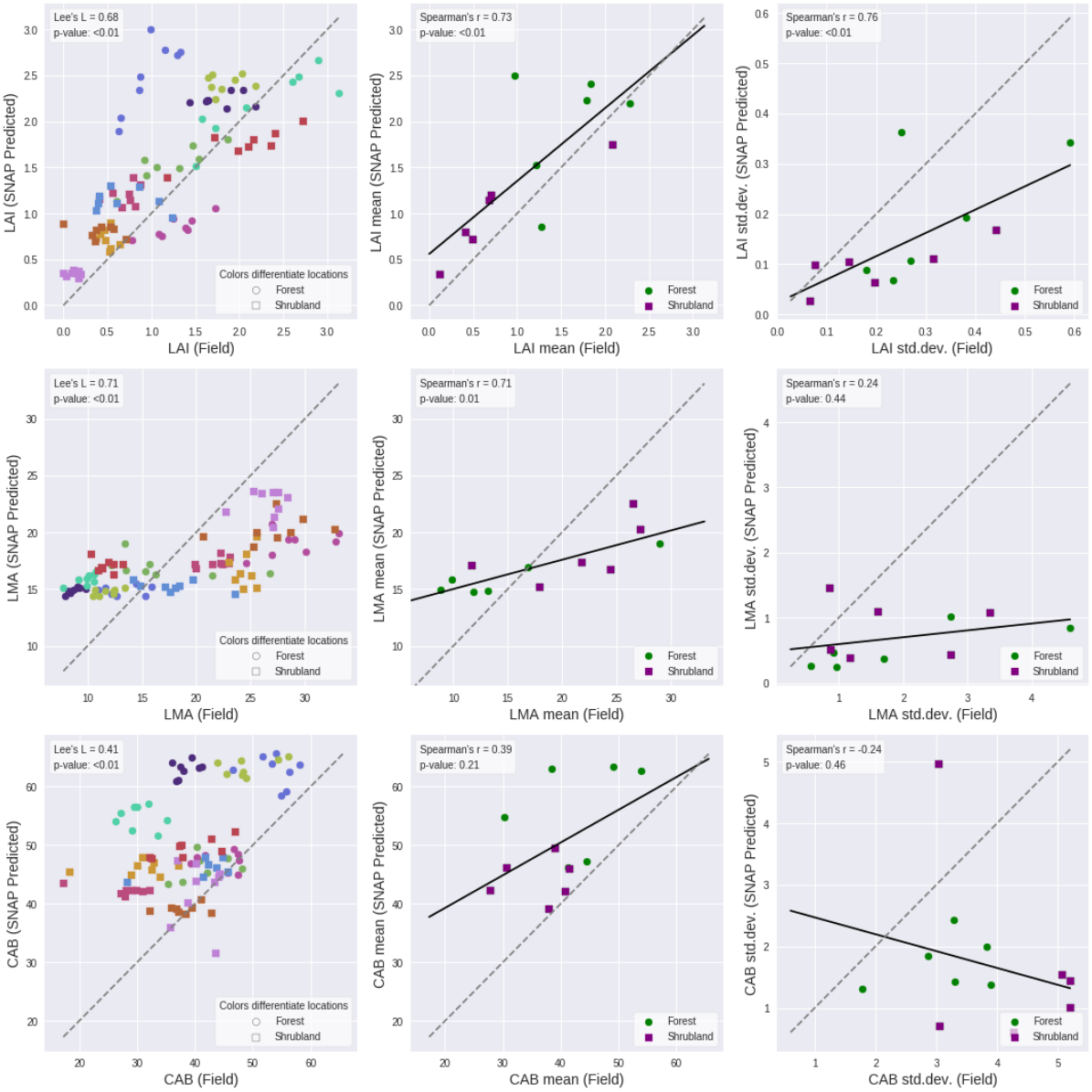


Fig. S. 12: Comparison of Sentinel-2-based trait estimates (y-axis) retrieved using the SNAP biophysical processor against in-situ field measurements (x-axis). The left column shows pixel-level (in-situ: plot) comparisons of traits, where different colors indicate plots of respective locations. The middle column depicts trait means per location and the right column presents trait standard deviations per location. The grey dotted line shows the 1:1 relationship, whereas the black line indicates the fitted linear relationship between the remotely sensed estimates and field data. Purple and green markers represent shrubland and forested locations, respectively.

S. 13: SNAP biophysical estimates of functional diversity compared against in-situ observations of functional diversity and taxonomic diversity

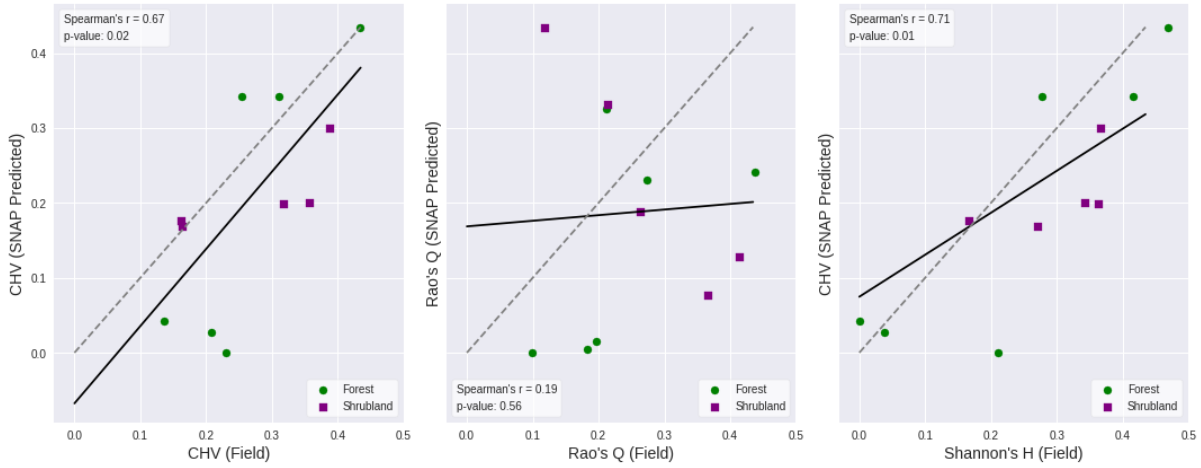


Fig. S. 13: Left and center panels present remotely sensed functional diversity estimates (CHV and Rao's Q) calculated from Sentinel-2 derived traits (LAI, LMA, and CAB) by the SNAP biophysical processor compared against in-situ functional diversity measurements. The rightmost plot compares SNAP-derived remotely sensed functional diversity (CHV) against in-situ taxonomic diversity (Shannon's H). The grey dotted line shows the 1:1 relationship, whereas the black line indicates the linear relationship between the remotely sensed estimates and field data. Purple and green markers represent shrubland and forested locations, respectively.

S. 14: In-situ correlations between functional diversity (CHV and Rao's Q) and community taxonomic diversity (Shannon' H)

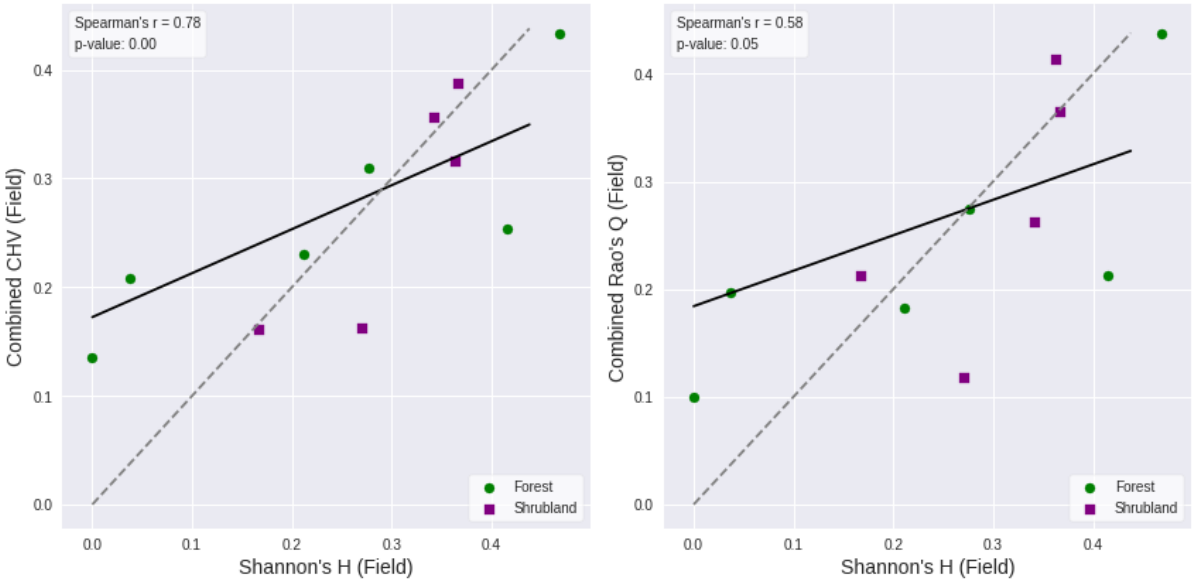


Fig. S. 14: Correlations of in-situ observations of functional diversity, CVH (left column) and Rao's Q (right column), with taxonomic diversity Shannon's H) as observed on the ground. Purple and green markers represent shrubland and forested locations, respectively.

S. 15: Alternative scaling (leaf versus canopy) of traits for functional diversity estimates retrieved by SVR hybrid inversion.

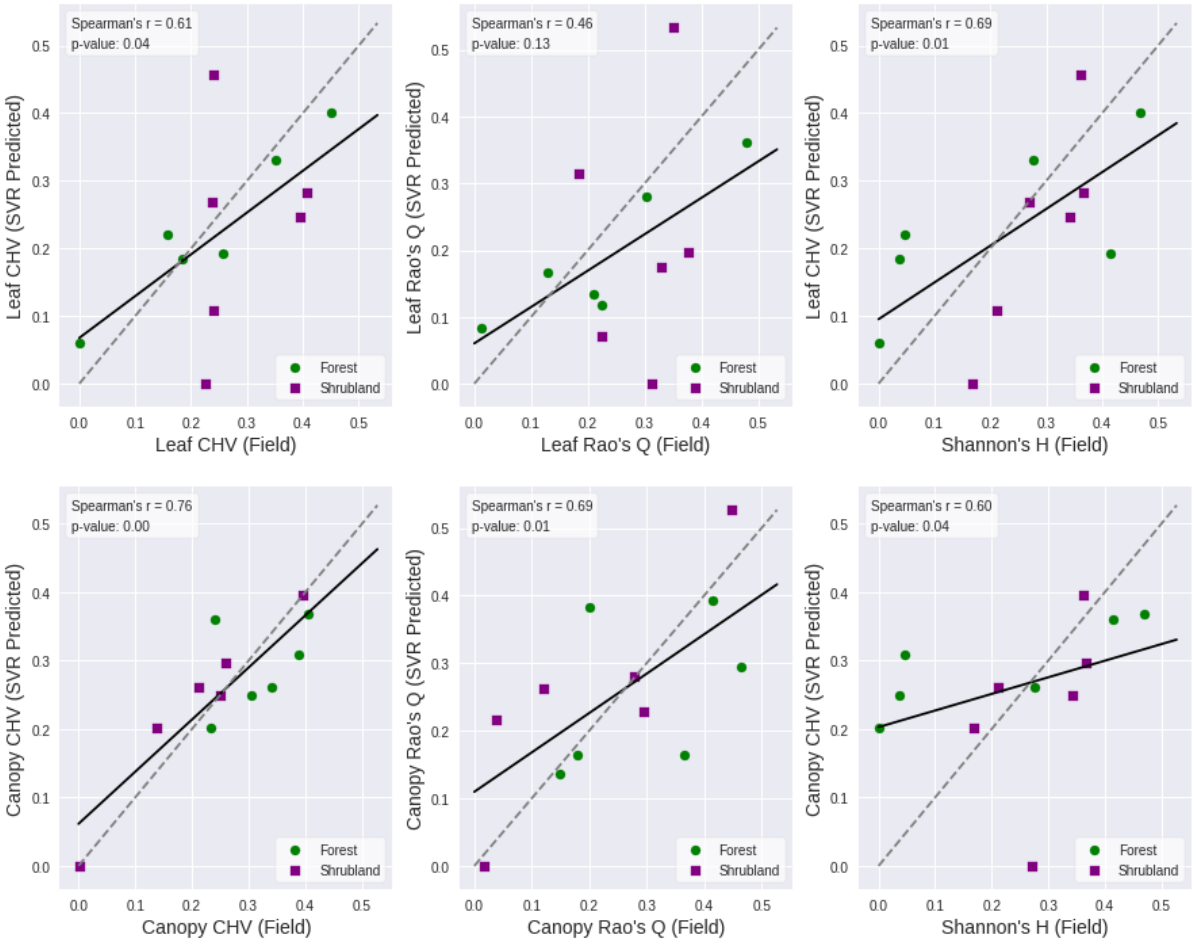


*Fig. S. 15: Remotely sensed functional diversity estimates (CHV and Rao's Q) calculated from Sentinel-2 derived leaf-level (top row) versus canopy-level (bottom row) scaling of traits through SVR hybrid inversion. The top row compares estimated functional diversity based on leaf-level traits (LMA and CAB) against in-situ observations of functional diversity and taxonomic diversity. The bottom row does the same comparing estimated functional diversity based on canopy-level traits (LMA*LAI and CAB*LAI) against in-situ observations. The grey dotted line shows the 1:1 relationship, whereas the black line indicates the linear relationship between the remotely sensed estimates and field data. Purple and green markers represent shrubland and forested locations, respectively.*

S. 16: Alternative scaling (leaf versus canopy) of traits for functional diversity estimates retrieved by ARTMO LUT-based inversion.

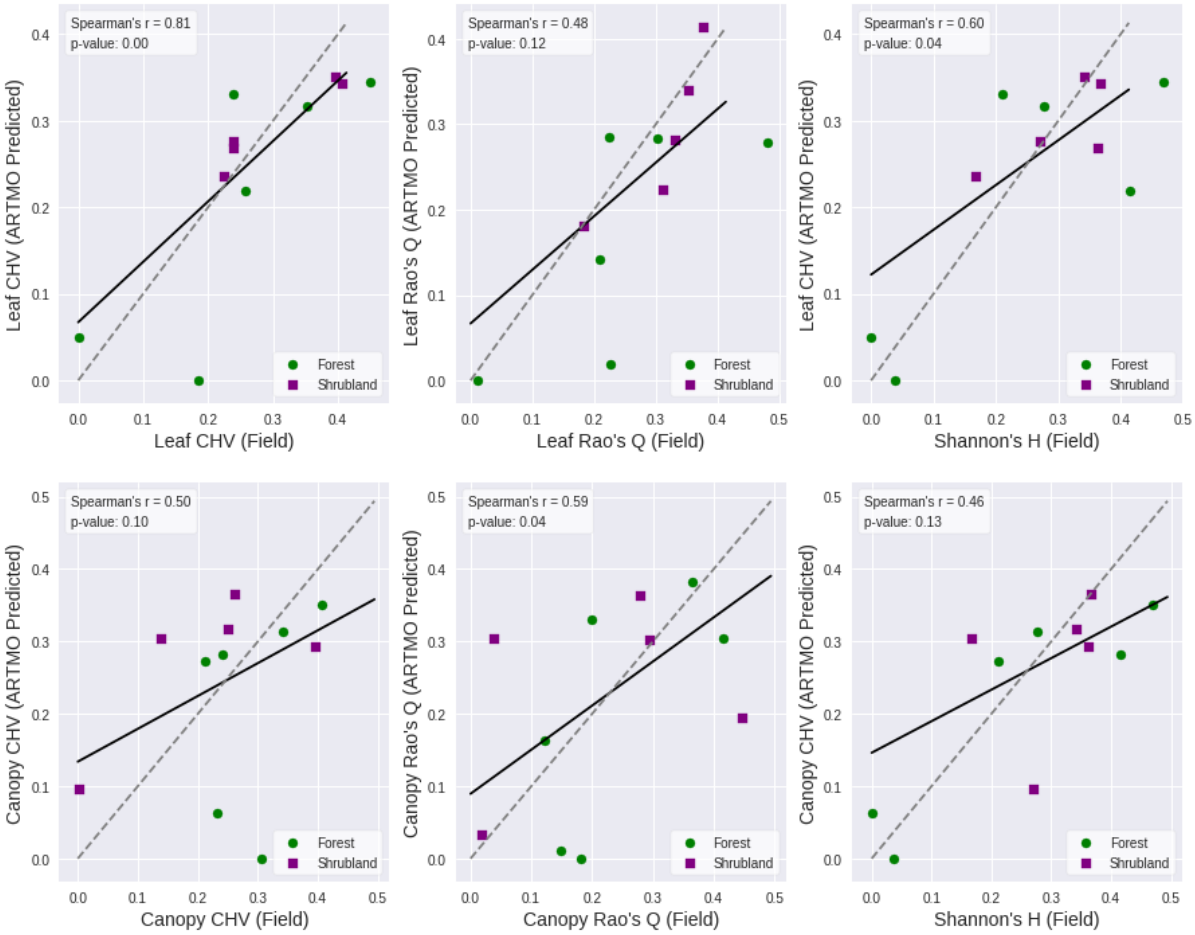


*Fig. S. 16: Remotely sensed functional diversity estimates (CHV and Rao's Q) calculated from Sentinel-2 derived leaf-level (top row) versus canopy-level (bottom row) scaling of traits through ARTMO LUT-based inversion. The top row compares estimated functional diversity based on leaf-level traits (LMA and CAB) against in-situ observations of functional diversity and taxonomic diversity. The bottom row does the same comparing estimated functional diversity based on canopy-level traits (LMA*LAI and CAB*LAI) against in-situ observations. The grey dotted line shows the 1:1 relationship, whereas the black line indicates the linear relationship between the remotely sensed estimates and field data. Purple and green markers represent shrubland and forested locations, respectively.*

S. 17. Alternative scaling (leaf versus canopy) of traits for functional diversity estimates retrieved by the SNAP biophysical processor.

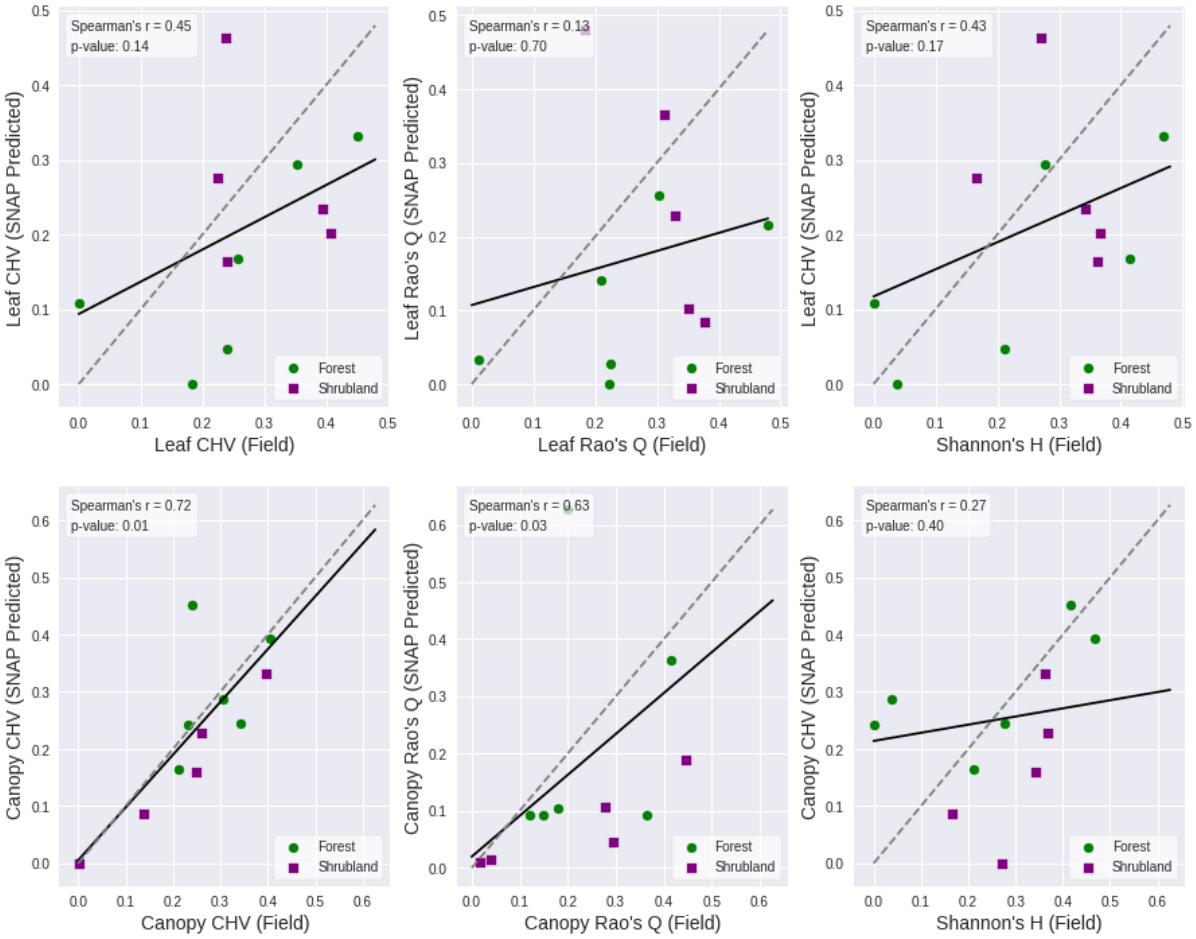


*Fig. S. 17: Remotely sensed functional diversity estimates (CHV and Rao's Q) calculated from Sentinel-2 derived leaf-level (top row) versus canopy-level (bottom row) scaling of traits through the SNAP biophysical processor. The top row compares estimated functional diversity based on leaf-level traits (LMA and CAB) against in-situ observations of functional diversity and taxonomic diversity. The bottom row does the same comparing estimated functional diversity based on canopy-level traits (LMA*LAI and CAB*LAI) against in-situ observations. The grey dotted line shows the 1:1 relationship, whereas the black line indicates the linear relationship between the remotely sensed estimates and field data. Purple and green markers represent shrubland and forested locations, respectively.*

S. 18: Alternative scaling (leaf versus canopy) of traits for in-situ correlations between functional diversity (CHV and Rao's Q) and community taxonomic diversity (Shannon' H).

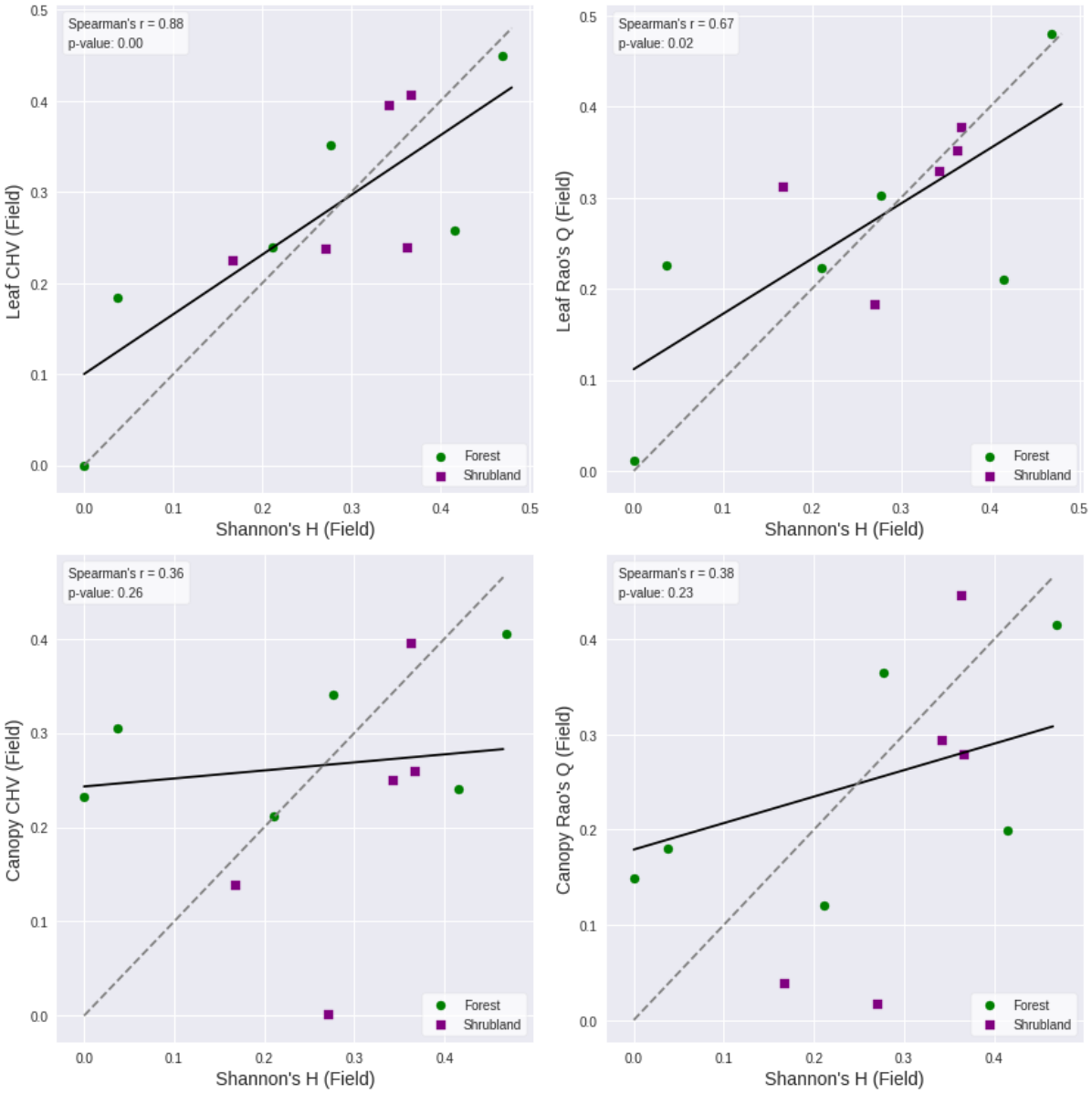


*Fig. S. 18: Correlations of CVH (left column) and Rao's Q (right column) with taxonomic diversity Shannon's H) as observed in-situ. The top row compares in-situ observations of functional diversity based on leaf-level traits (LMA and CAB) against in-situ observations of taxonomic diversity. The bottom row does the same comparing canopy-level traits (LMA*LAI and CAB*LAI) in-situ functional diversity against in-situ taxonomic diversity. The grey dotted line shows the 1:1 relationship, whereas the black line indicates the linear relationship between the remotely sensed estimates and field data. Purple and green markers represent shrubland and forested locations, respectively.*

Fig. S. 19: a) Map of the Sentinel-2 observation, as seen through a true visible range composite image, over the study area in Sabah, Malaysia on the 9th of July 2017 after atmospheric corrections. b) Map of the three land use classes within the Sabah study area used for analysis.

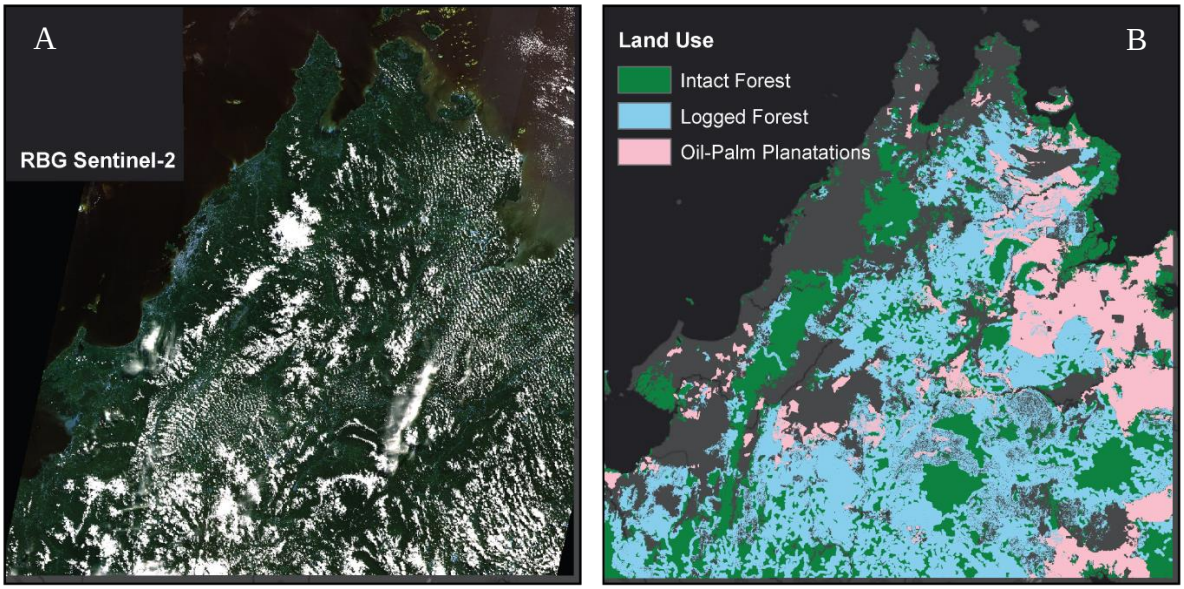


Table S. 20: Descriptive statistics of elevation and slope observed between plots across land use types. Data is derived from digital elevation models observed by the 30m spatial resolution Shuttle Radar Topography Mission (SRTM).

	N	Mean elevation (m)	Mean slope (% change)
Intact Forest	1192	439.90	1.88
Logged Forest	3018	419.71	1.30
Oil-Palm Plantations	1416	63.84	0.74

S. 21: Qualitative assessment of the validity of spectral trait indicators

To assess the plausibility of the derived spectral trait indicators on which we based functional diversity estimates, the performance of the inversion of SNAP's biophysical processor retrieval from spectra to traits was examined.

Firstly, we conducted a sensitivity analysis to assess the spectral layout of Sentinel-2 bands in terms of receptiveness to retrieve these traits (Fig. S. 21a). The analysis is based on repeated PROSAIL simulations with random variations of the trait values while mapping the spectral responses and the correlation between Sentinel-2's spectral bands and trait variation. The findings demonstrate significant spectral sensitivity to the studied traits (LAI, EWT/LWC, and CAB) (Fig. S. 21a). These findings match the sensitivity analyses of Sentinel-2 for retrieval of the traits under study (LAI, CAB and EWT/LWC) demonstrated in previous sensitivity analyses (de Sá et al., 2021; Gu et al., 2016; Rossi et al., 2020; Verrelst et al., 2019b).

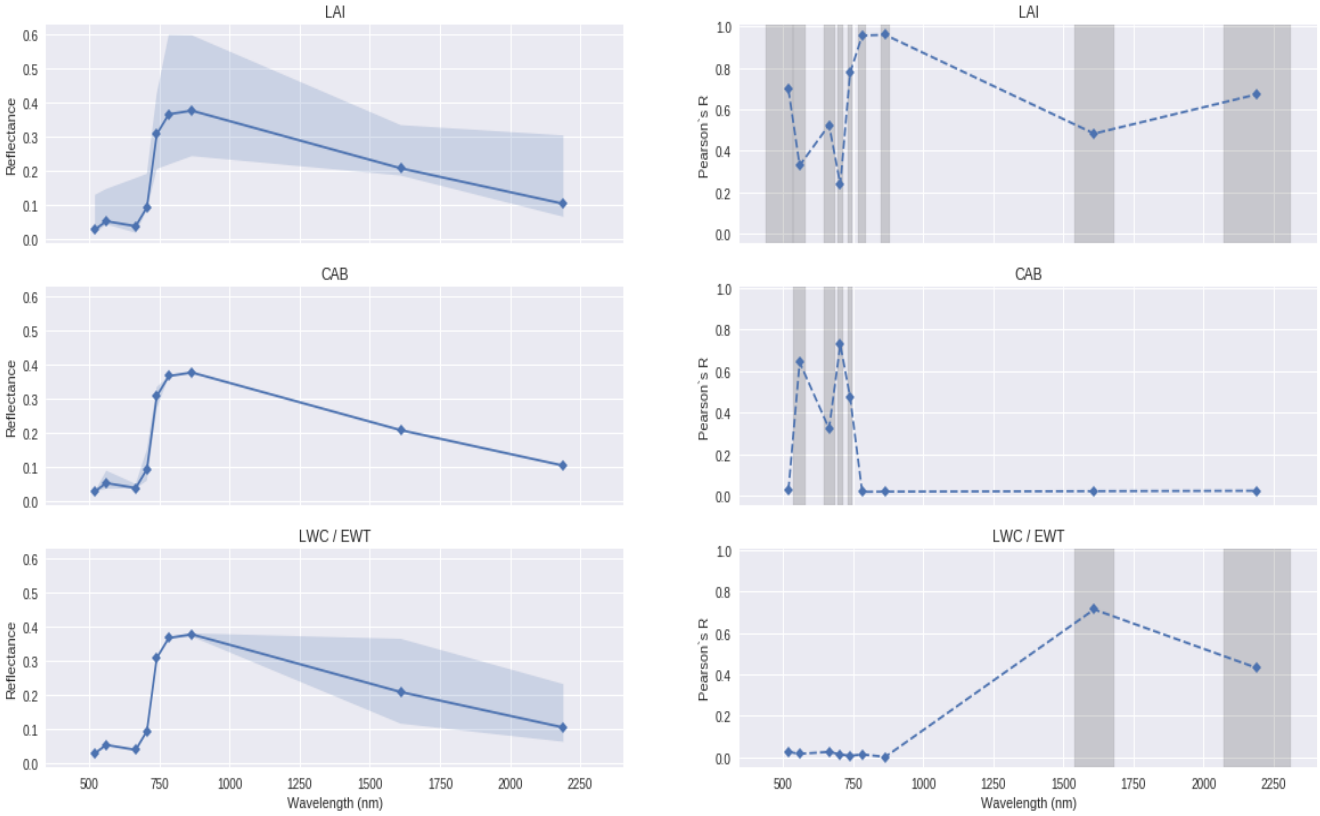


Fig. S. 21a: Sensitivity analysis of PROSAIL's LAI, CAB and EWT/LWC parameters to the spectral lay-out of Sentinel-2 MSI bands used in this study. The left pane illustrates the range of variability in spectral response to changes in the parameters while all other parameters are kept constant at modes defined in Weiss & Baret (2016). The right pane depicts the correlation (Pearson's R) of different bands to changes the parameters using a range of simulations defined in Weiss & Baret (2016). Significant correlations are indicated by the grey highlighted bars.

Secondly, the presence of biases in the derived spectral trait indicators across land use types was examined by reversing the inversion process. From the spectrally retrieved trait estimates obtained through inversion, we re-modeled spectra based on PROSAIL in forward-mode. The simulated spectra were compared against the actual observed Sentinel-2 spectra to assess the

performance (size of the error) between different land use types/canopy types. This was done for 20,000 randomly drawn pixels over the study area. Search ranges for the remaining PROSAIL traits were constrained to the distribution of input variables described by Weiss & Baret (2016). We used Spectral Angle Mapper (SAM), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) to assess the deviation between the simulated and actual spectra. We evaluated differences in the size of errors across land use types to examine whether such biases could have affected our results. On average, across the land use types, a RMSE of 0.012 is observed which is a 7.6 percent mean deviation from the actual observed spectra (See table S3.2). No profound differences in errors were observed between the individual land use types which suggest there is no structural bias in the inversion performance for approximation of the vegetation present in different land use.

Table S. 21b: Comparison of errors between simulated spectra with retrieved trait estimates as input and actual Sentinel-2 reflectance spectra stratified across land use types.

	RMSE		MAE		%nRMSE	Sampled pixels (N)
Land Use	μ	σ	μ	σ	μ	
Intact Forest	0.012	0.009	0.007	0.005	7.6	4774
Logged Forest	0.014	0.01	0.008	0.006	8.9	4274
Oil-Palm Plantations	0.012	0.01	0.007	0.006	7.2	3952

Thirdly, we compared the distributions of spectral trait indicators against in-situ measured traits of common species in the different studied land use types. Figure S3.3 indicates that the spectral trait indicators are to a large degree in line with the range of measurements from field studies.

The most notable deviance is an overestimation in oil-palm plantations of LAI compared to Hadi et al. (2017)'s effective LAI measurements. Hadi et al. (2017)'s measurements are conducted in relatively recently planted oil-palm plantations. Differences in LAI between young plantations versus older plantations might be responsible for this discrepancy. In our study, observed LAI values –were the highest in plantations rather than tropical evergreen broadleaf forests. This was also reported by Asner et al. (2003). Intensive management regimes -through planting schemes, fertilizer, and pest management- aim to maximize production in these plantations and both resonate with relatively high chlorophyll and LAI ranges. Furthermore, early studies by Hardon et al. (1969) have demonstrated a positive correlation between the leaf area and the yield of different palms. Industrial oil-palm plantations are often located on more favorable lands for high productivity in terms of elevation and slope.

In terms of LWC/EWT and CAB, we find much larger variation in the species measurement conducted in the field. Notably, these measurements consist of a variety of individual species whereas our observations are based on pixel-based aggregation of multiple species possibly levelling out some of the variation. In consideration of this convergence, the range of spectral trait indicators is reasonable with the exception of an overestimation of LWC/EWT compared to Martin et al. (2018)'s in-situ measurements.

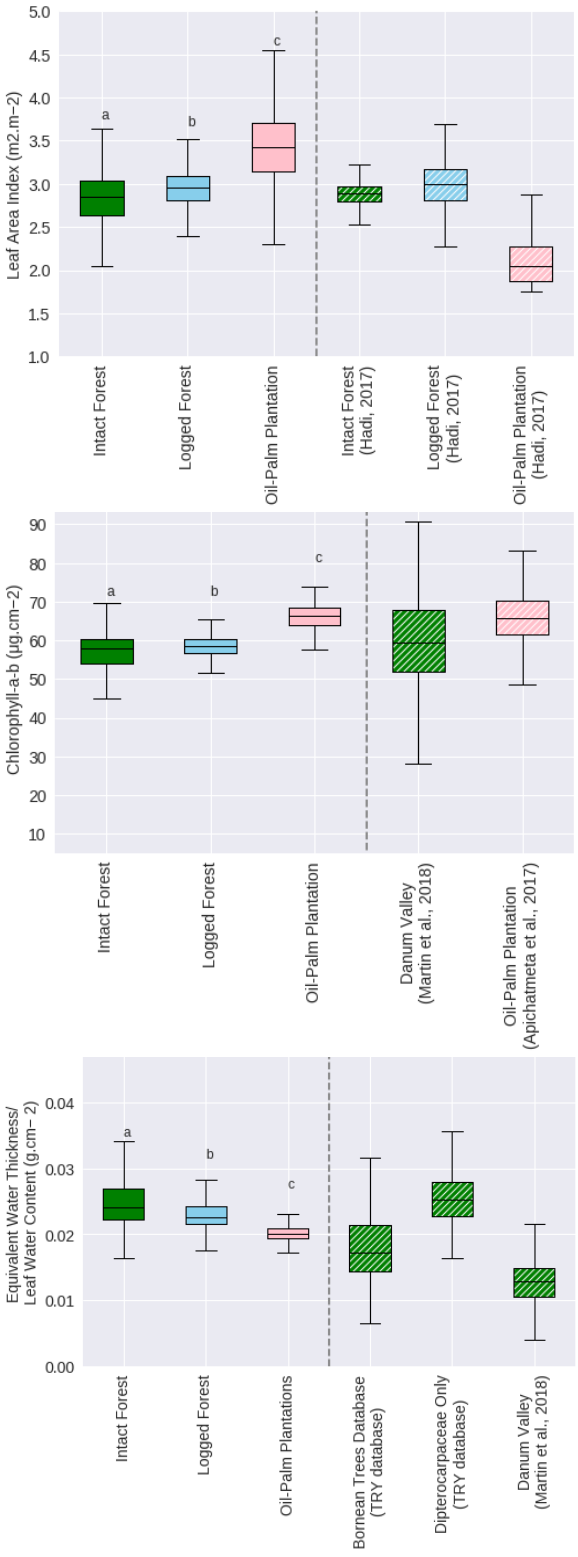


Fig. S. 21c: Comparison of retrieval of spectral trait indicators from SNAP's biophysical processor derived from Sentinel-2 reflectance spectra to relevant in-situ trait ranges in the TRY plant trait database (Kattge et al., 2011; Kurokawa and Nakashizuka, 2008) and other regionally relevant field campaigns (Apichatmeta et al., 2017; Azzeme et al., 2016; Hadi et al., 2017; Martin et al., 2018). Differences in retrieved trait estimates between land use types according to a post-hoc Tukey's HSD test are indicated with different letters.

S. 22: Functional diversity across elevation

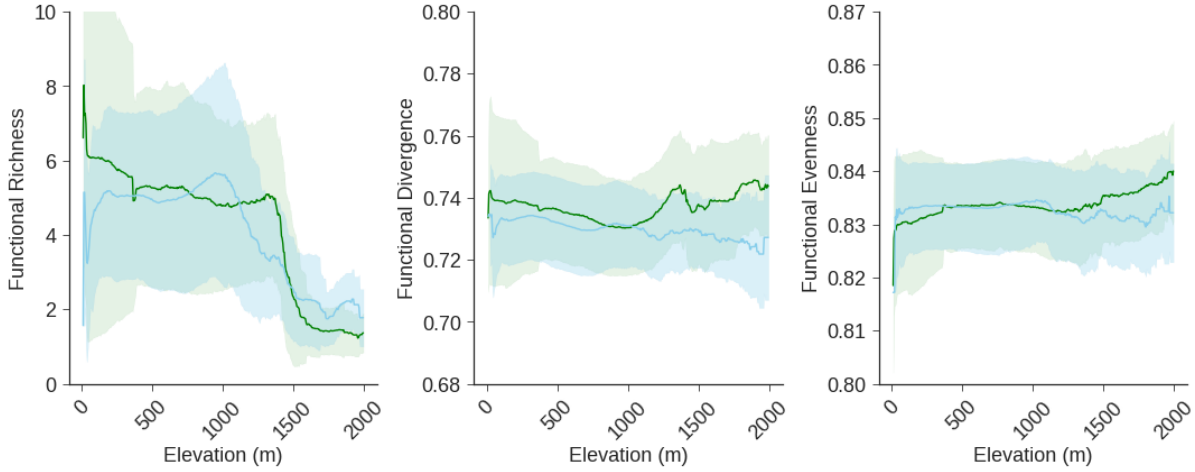


Fig. S. 22: Variation of functional diversity metrics along average elevation of plot. Green represents intact forests; blue indicates logged forests. Oil-palm plantations were only found below 500m ASL.

Functional richness decreases along the elevational gradient. In line with ecological theory, this indicates a stronger functional convergence with elevation linked to a stronger environmental filtering of fitness in higher altitudes (Durán et al., 2019). A steep drop in functional richness is particularly observable above 1400m ASL. These findings correspond with earlier studies on elevational patterns of tree species richness on Mount Kinabalu, Borneo (Aiba and Kitayama 2010; Grytnes and Beaman 2006; Kitayama 1992). Similar to functional richness in our study, tree species were found to decrease rapidly with elevation above 1500m (Grytnes and Beaman, 2006). Moreover, using airborne remote sensing, studies have shown a lower functional richness at higher elevation suggesting a smaller range of resource availability at higher altitudes, whereas, again, functional divergence remained relatively unaffected by elevation (Durán et al., 2019).

Generally, we observe a negative correlation of functional richness and elevation of Pearson's R = -0.18 for both land use types and of Pearson's R = -0.31 for Intact Forests specifically. For Logged Forests, we observe an initial increase in functional richness before following a similar decrease as seen in intact forests. Functional divergence and functional evenness, on the other hand, indicate little variation with elevation. Oil-palm plantations were not found to be grown at elevations above 500m above sea level (Table S. 20). This therefore did not allow for studying elevational effects in oil-palm plantations.

S. 23: Non-log-transformed functional-area relationships

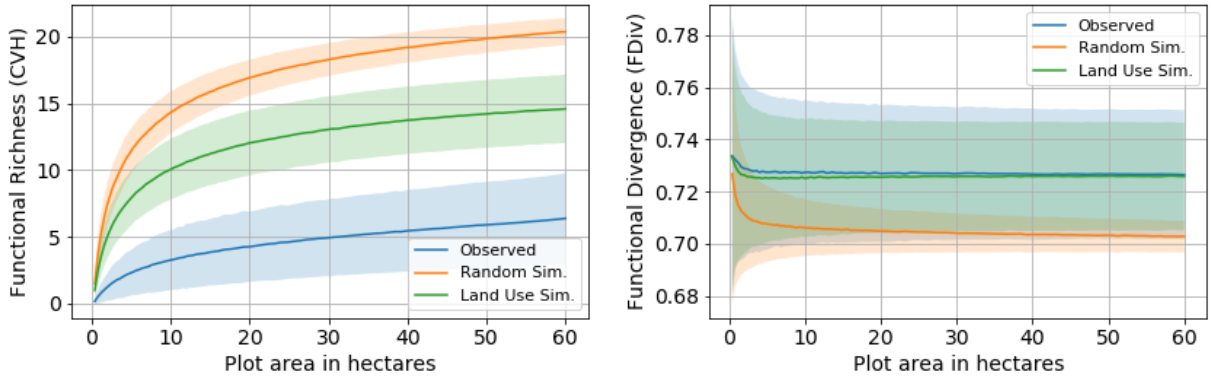


Fig. S. 23a: Non-log-transformed scale dependency of the two studied functional diversity metrics. Comparisons are made between observed remotely sensed diversity in Sabah, Borneo, versus null-models based on randomly drawn trait combinations from observed trait values. The 'Random Sim.' model (orange) draws random trait combinations irrespective of land use type, whereas the 'Land Use Sim.' model draws random trait combination in consideration of land use. Dark lines represent the mean and standard deviations are represented by the respective shaded areas.

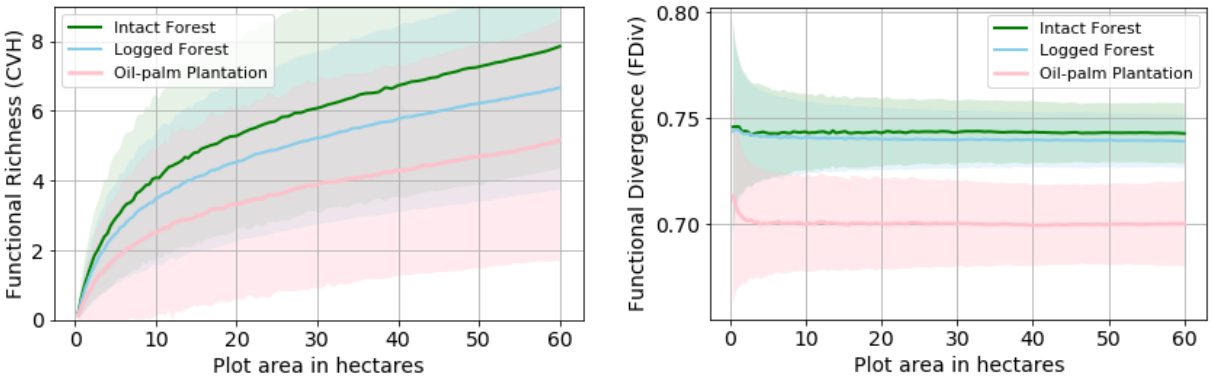


Fig. S. 23b: Non-log-transformed scale dependency of the two studied functional diversity metrics stratified across land use types. Functional diversity calculations were done over continuous plots of one land use type over a range from 0.4 to 60 ha sized plots. Dark lines represent the mean and standard deviations are represented by the respective shaded areas.

S. 24: Posthoc analysis of functional diversity differences between land use types

S. 24: Posthoc analysis of ANOVA to assess the significant differences between land use pairings. P-values below 0.05 are considered to signal significant differences between individual land use types.