Title
Mapping multiple dimensions of forest diversity using spaceborne spectroscopy

Authors
J. Antonio Guzmán Q.[1,2]*, Jonathan A. Knott [3], Jesús N. Pinto-Ledezma [2], Philip A. Townsend [4], Jeannine Cavender-Bares [1,2]*

Affiliations
[1] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA.
[2] Department of Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, MN 55108, USA.
[3] USDA Forest Service, Northern Research Station, Saint Paul, MN 55108, USA.
[4] Department of Forest and Wildlife Ecology, University of Wisconsin—Madison, Madison, WI 53706, USA.

*Corresponding authors: aguzman@fas.harvard.edu and jcavender@fas.harvard.edu

Abstract
Observing biodiversity across space and time is essential for advancing and verifying conservation efforts toward global biodiversity and sustainability goals. Spaceborne imaging spectroscopy has emerged as a revolutionary tool for quantifying and tracking forest diversity, yet its application at large spatial scales remains a central challenge. We develop a framework to map multiple dimensions of forest community composition and diversity by integrating imaging spectroscopy from two spaceborne sensors (DESIS and EMIT) with taxonomic, phylogenetic, and functional trait datasets, and 43,155 forest inventory plots across the Eastern United States. Our findings show that satellite-based spectral dissimilarity among forest communities is positively correlated with estimations of β-diversity from ground inventories. We further demonstrate that imaging spectroscopy can be used to predict ordination axes of β-diversity, enabling the mapping of multiple dimensions of forest community composition over large spatial extents. We showcase how these β-diversity ordinations and maps support the distribution modeling of 95 forest attributes and allow the evaluation of spatiotemporal changes in community composition. Our framework demonstrates that spaceborne imaging spectroscopy, when combined with inventory data, allows indirect yet comprehensive observation of forest diversity. This integrative approach sets the stage for scalable forest monitoring in support of global biodiversity conservation and forthcoming satellite missions.

Keywords: Biodiversity, β-diversity, DESIS, EMIT, Forest inventories, Hyperspectral observations.

Highlights
- Hyperspectral data and inventories were integrated to map forest diversity broadly
- Satellite spectral dissimilarity correlates with ground-based β-diversity
- DESIS and EMIT enable large-scale mapping of forest community composition
- Mapped forest composition supports modeling of forest types, lineages, and traits
- Mapped forest composition enables detection of forest community change over time

1. Introduction

Monitoring biodiversity and its changes across space and time represents a major scientific and societal challenge critical for sustainable management of our planet (Cavender-Bares et al., 2022; Gonzalez et al., 2023). The accelerating loss of biodiversity and shifts in species composition resulting from human pressures and global change underscore the urgent need to develop effective approaches for biodiversity monitoring (IPBES, 2019). Reliable and scalable observations are essential not only for verifying efforts towards the sustainability targets of the Convention on Biological Diversity and the Sustainability Development Goals, but also for informing ecological forecasting and guiding adaptive management strategies (Leclère et al., 2020). Biodiversity observation and monitoring are particularly important in forest ecosystems, which provide essential ecosystem services and are highly vulnerable to land use change, pests, pathogens, and abiotic stressors. Forest biomes in North America represent a high-priority region for monitoring given their vast extent and critical role in supporting ecosystem services and the economy (Cavender-Bares et al., 2022).

Our capacity to observe biodiversity across broad spatial scales is hindered by the inherent spatial and sampling limitations of traditional field surveys. Furthermore, field surveys are also constrained by financial and logistical challenges, including inaccessibility of remote areas or private lands (Jetz et al., 2016; Westfall et al., 2022). As a result, imaging spectroscopy has been proposed as a transformative approach for observing biodiversity and advancing its conservation (Turner, 2014). In particular, spaceborne spectroscopy offers a promising avenue for mapping forest diversity due to its global coverage and spatial and spectral resolutions sufficient for biodiversity assessment (Jetz et al., 2016; Turner et al., 2015). Nonetheless, many current efforts to map plant diversity have relied on airborne or near-surface platforms (Draper et al., 2019; Féret and Asner, 2014; Gholizadeh et al., 2020, 2019; Pinto-Ledezma et al., 2025; Rossi et al., 2022; Schweiger et al., 2018; Schweiger and Laliberté, 2022; Wallis et al., 2024; Wang et al., 2018). Compared to these platforms, spaceborne spectroscopy typically offers coarser spatial resolution (e.g., 30–60 meters), which tends to capture forest communities rather than individual trees or crowns. In addition, airborne and near-surface approaches often cover limited regions and ecosystems due to their restricted spatial extent. Consequently, there is a clear need for spaceborne spectroscopy to fill the gap of continuous forest diversity observations over extensive areas (Skidmore et al., 2021). The capability of using satellite spectroscopic imagery to map multiple dimensions of forest diversity remains a central challenge.

A growing body of literature demonstrates that spatial dissimilarity in species composition among communities—known as taxonomic beta (β)-diversity—is positively associated with spectral dissimilarity (Dalmayne et al., 2013; Everest et al., 2025; Schweiger and Laliberté, 2022). This correspondence between compositional and spectral dissimilarity has been observed using both hyperspectral airborne data and vegetation indices derived from satellite imagery (Dalmayne et al., 2013; Everest et al., 2025; Schweiger and Laliberté, 2022). Likewise, spectral dissimilarity among species has been shown to correlate with phylogenetic and functional differences, with more distantly related or functionally distinct species exhibiting greater spectral divergence (Pinto-Ledezma et al., 2025; Schweiger et al., 2018; Wallis et al., 2024). These findings suggest that integrating spectral dissimilarity with measures of community dissimilarity can enable the remote observation of forest diversity at broad spatial scales, particularly in regions with well-characterized ground-based data. Community dissimilarity metrics predicted from satellite imagery have the potential to reveal multiple dimensions of forest diversity and to enable detection of species turnover and nestedness by partitioning β-diversity into its replacement and richness-difference components (Baselga, 2010; Cardoso et al., 2014; Pinto-Ledezma et al., 2018).

Here, we present a novel framework for comprehensive mapping of forest community composition, including taxonomic, phylogenetic, and functional dimensions, at a large spatial extent using spaceborne spectroscopy (Fig. 1). We integrate ground observations from the USDA Forest Service Forest Inventory and Analysis (FIA) Program with spaceborne spectroscopy data from the DLR Earth Sensing Imaging Spectrometer (DESIS) and the Earth Surface Mineral Dust Source Investigation (EMIT) across the Eastern United States. To do so, we first evaluate the relationship between multiple dimensions of β-diversity and spectral dissimilarity among forest communities. Our general hypothesis is that forest communities with greater dissimilarity in taxonomic, phylogenetic, or functional composition exhibit higher spectral dissimilarity than communities with similar compositions. We test this hypothesis and then assess the extent to which spectroscopy observations can predict β-diversity ordinations to provide biodiversity indicators for mapping dimensions of forest community composition. Finally, we present potential applications of β-diversity ordinations and community composition maps within the context of distribution models. These applications include prediction of forest types, plant lineages, and community-weighted means (CWM) of plant traits, as well as evaluation of spatiotemporal changes in forest community composition. This study provides a novel approach for mapping forest diversity and its multiple dimensions over large spatial extents by leveraging the intrinsic relationships between forest community composition and its optical properties as captured by spaceborne spectroscopy.

2. Materials and Methods

2.1 Study area

Our study was conducted in the Eastern continental U.S. region with an area close to 2,593,107 km^2 (Fig. S1). This region encompasses eight eco-climatic domains defined by the National Ecological Observatory Network (NEON) (Hargrove and Hoffman, 1999), each with distinct vegetation, climate, and ecosystem dynamics, which were used to delineate our study area. Despite its extension, this region presents a comparable timing in the peak of greenness (i.e., June through August) (Bolton et al., 2020; Brooks et al., 2020), which can be used to restrict the phenological effects on the optical properties of forest communities.

2.2 Forest inventory data

We used inventory data collected from the Forest Inventory and Analysis (FIA) program of the United States Department of Agriculture Forest Service (USDA Forest Service, 2023). We used data from 43,155 FIA plots across the Eastern United States (Fig. S1), encompassing 15,550,151 individual trees representing 243 species, 45 genera, and 23 families. These forested plots were inventoried between January 1st, 2018, and January 12th, 2023. Each inventory plot consists of a central circular subplot (7.31 m radius) surrounded by three circular subplots at 36.57 m from their centroids located at 0-, 120-, and 240-degree azimuths (Bechtold and Patterson, 2015). Trees (> 12.7 cm diameter at breast height) from all the subplots were used to characterize local tree communities given the spatial resolution of the spaceborne observations. We estimate wood volume for each live tree assuming a cylinder shape using the diameter at breast height and tree height and then expand its value to per-unit-area using the FIA trees-per-acre expansion factor. The resulting wood volume per unit of area was then summed per species as a descriptor of species abundance for further analysis. We used the forest type with the largest proportion of occurrence that is recorded on each FIA plot as a descriptor of the forest community for further analysis. More details about the selection of inventories and filtering of plots are shown in Methods S1.

2.3 Taxonomic, phylogenetic, and functional β-diversity

Among each plot we estimated β total diversity (β_{total}) following improvements by (Cardoso et al., 2014) based on pioneering work by (Baselga, 2010) as a framework to estimate the taxonomic, phylogenetic, and functional dissimilarities among tree communities. Within this framework, pairwise comparisons of communities (i.e., plots) were performed to partition β-diversity into components of replacement ($\beta_{replacement}$, i.e., diversity explained by replacement of species alone) and richness ($\beta_{richness}$, i.e., diversity explained by species loss/gain alone). These were used to compute β total diversity (β_{total}) as the sum of $\beta_{replacement}$ and $\beta_{richness}$ (See more details in (Cardoso et al., 2014)). We used Jaccard dissimilarity weighted by the relative abundance of the wood volume per unit of area in all the estimations of β. Specifically, we used a pruned phylogenetic tree from (Smith and Brown, 2018) to estimate phylogenetic β (Pβ) (Fig. S2). The phylogenetic tree was obtained through the 'phylo.maker' function of the V.PhyloMaker2 package (Jin and Qian, 2022) in R (R Core Team, 2023) using its third scenario. Under the third scenario, V.PhyloMaker2 adds missing species to the half point of their parent branch and uses the BLADJ approach for branch length estimation. For functional β diversity (Fβ), we used eight plant traits for each species (Methods S2 and Fig. S2) to summarize their variation into three principal components (Fig. S4) and construct a dendrogram forced as a phylogeny. Taxonomic β diversity (Tβ) was estimated using species taxonomic names occurring in the plots. β values were estimated using a modification of the BAT package by (Cardoso et al., 2015) with C^{++} and OpenMP as a backend to efficiently run a large number of pairwise comparisons between communities in parallel using high-performance computing.

2.4 Ordinations of β-diversity

To reduce the dimensionality of the taxonomic, phylogenetic and functional trait matrices created from the forest inventory data, we applied landmark Multi-Dimensional Scaling (MDS) to ordinate communities in three-dimensions (i.e., axes), according to their β_{total} for each dimension of forest diversity. Each MDS was rotated to match mean annual temperature and annual precipitation gradients derived from Worldclim 2 (Fick and Hijmans, 2017), using a sequential orthogonal rotation procedure that aligns the ordination axes with the fitted environmental vectors via vector fitting and planar rotations. To evaluate each ordination and their meaning, we first calculated the goodness of fit as the proportion of variance (R^2) of β_{total} that is accounted for by euclidean distances in MDS β-diversity axes values between communities. Then, climatic, topographic, community weighted mean (CWM) of plant traits, and the relative abundance of major plant lineages (i.e., gymnosperms, angiosperms, arbuscular mycorrhizal-, ectomycorrhizal-symbiosis trees) were projected into these ordinations by computing the weighted means based on the axes using the 'wascores' function of *vegan* (Oksanen et al., 2025). We used MDS β-diversity axes to model dimensions of forest community composition from spaceborne observations, and to estimate then community plant traits and the occurrence probability forest types and plant lineages.

2.5 Spaceborne hyperspectral data and processing

We used scenes from two imaging spectrometers that are docked to the International Space Station: the DLR Earth Sensing Imaging Spectrometer (DESIS) (i.e., 30 m pixel resolution) and the Earth Surface Mineral Dust Source Investigation (EMIT) (i.e., 60 m pixel resolution). For both sensors we used Level 2A hyperspectral scenes with less than 70% of cloud cover that provide surface reflectance data corrected by different means (e.g., *46*, *47*). For DESIS, we only employed scenes (n = 3145) within the peak of the growing season between June 15th to August

15th from 2019 to 2023. For EMIT, on the other hand, we only used scenes (n = 349) collected between June 1st to August 30th of 2023 and 2024. The spectral reflectance from both sensors were transformed using continuous wavelet transformation (CWT) in order to enhance absorption features, reduce angular/illumination effects, and smooth the spectral signal (Rivard et al., 2008). For this, we first resample the reflectance spectra to a continuum band spacing (DESIS: 3.00 nm; EMIT: 7.43 nm) using the Full-Width-Half-Maximum method. CWT was applied then on the resampled spectra by selecting different scales (DESIS: 2^3, 2^4, and 2^5; EMIT: 2^2 and 2^3) that capture the combination of small and large reflectance features. These scales were summed to create a summed-wavelet spectra. Bands from the transformed spectra with potential noise, atmospheric contamination, or close to the edge of the spectral range of the sensors were removed. We ended with scenes of 173 bands with a spectral range between 449 – 965 nm for DESIS, while scenes of 179 bands with spectral ranges between 448 – 1236, 1521 – 1686, and 2053 – 2421 nm for EMIT. Our analyses excluded pixels with Normalized Difference Vegetation Index (NDVI, $\sim\lambda_{800}$ - $\sim\lambda_{680}$ / $\sim\lambda_{800}$ + $\sim\lambda_{680}$) and NIR values ($\sim\lambda_{800}$ nm) lower than 0.4 and 0.3, respectively, to ensure that observations come from vegetated surfaces and exclude shadows, respectively. Using federal-protected coordinates FIA plots we extracted 11,526 and 10,469 clear-sky pixels from DESIS and EMIT imagery, respectively, corresponding to locations with spatial overlap of scenes with forest inventory data.

2.6 Relationship between β-diversity and spectral dissimilarity

We estimate spectral dissimilarity between plots in the inventory data using the Spectral Angle Mapper (SAM) method (Kruse et al., 1993). Using SAM, reflectance spectra are n-dimensional vectors (i.e., number of bands) to estimate the spectral angle (i.e., 0° – 90°) between two communities and then estimate a metric of dissimilarity (i.e., 0 – 1), where values close to 1 describe forest communities with contrasting spectra. Using each inventory as a reference, we computed pairwise matrices of spectral dissimilarity between all plots (Methods S3). To evaluate the influence of spectral range and spatial resolution, we calculated spectral dissimilarity for EMIT using both its full spectral range and the spectral range shared with DESIS; DESIS and EMIT were then compared using this common spectral region. We performed Mantel tests to assess the association between matrices of spectral dissimilarity and dimensions of β and their partitions using Pearson correlations and 999 permutations. Then, using each community as a reference, we fitted separate linear models of spectral dissimilarity against multiple dimensions of β and their partition (i.e., β_{total}, $\beta_{replacement}$, $\beta_{richness}$), based on pairwise comparisons with all other communities.

2.7 Modeling multiple dimensions of β-diversity

We predicted multiple dimensions of β-diversity using FIA plots that overlap with the available transformed hyperspectral scenes. For this, we extracted pixels to predict MDS scores for each dimension of β_{total} using a modeling framework based on partial-least squares regression (PLSR). The extraction of pixels was done using the federally protected locations of FIA plots obtained from the National Information Management System internal to the FIA program. FIA plots may have multiple recorded plot locations from different remeasurement years (with GPS errors up to 10 m, (Knott et al., 2023)), so to link to the imagery, we calculated the mean latitude and longitude from all remeasurements (when available) of each individual FIA plot following (McRoberts et al., 2018). From DESIS scenes, we restrict the pixel extraction of overlapping inventories with less than ± 6 years of difference between the inventory and observation date. For each plot, we limit the number of extracted scene observations by selecting the clear-sky observation closest to the inventory date. For further analysis we employed a total of 11 526 and 10 469 pixels from DESIS and EMIT scenes, respectively (Fig. S1).

Our modeling framework consisted of repeated (n = 100) PLSRs using a spatial cross-validation strategy based on stratified random sampling. For this, we first randomly split 60% of the samples available per county for algorithm training to ensure a balanced representation of forest inventories across counties. Then we applied spatial cross-validation models using a subset of training samples on each iteration. For each iteration and axis, we selected 85% of training samples, sampling randomly across histogram distributions to capture the range of axis variation along with nominal breaks. We used a 10-fold leave-locations-out cross-validation where a series of samples from counties are spatially excluded to train the model. This repeated framework helps us to compute the mean and SD of the model estimates, and thus the potential uncertainty of the predictions. Once the repeated models were computed, we determined the optimal number of components based on the spatial cross-validation error of the RMSE of prediction (Kuhn and Johnson, 2013; Meyer et al., 2018). With the optimal number of components, we then estimate the model performance on both training and testing datasets by computing the coefficient of determination (R^2), the bias, the root mean squared error (RMSE), and the percentage RMSE (%RMSE = RMSE/ range of 0.99 and 0.01 quantiles × 100). From these repeated models, we extracted variable importance in projection (VIP) scores to assess which spectral regions were most important for predicting MDS β-diversity axes, and PLSR coefficients to apply to spaceborne imagery for mapping forest community composition.

2.8 Modeling forest attributes

We predicted the occurrence of forest types and plant lineages, and the variation of community weighted means (CWM) of plant traits using MDS β-diversity axes, as well as climatic and topographic variables (Methods S4). To generate these models, we followed an approach similar to (Gu et al., 2015), but we focused on occurrence probabilities and CWM of plant traits. Specifically, we used $T\beta_{total}$-base MDS ordinations to predict true presence or absence of forest types, $P\beta_{total}$-base MDS ordinations to predict the true presence or absence of a plant lineage (e.g., plant family, plant genus, or mycorrhizal symbiosis) within a community, and $F\beta_{total}$-base MDS ordinations to predict the CWM of plant traits. Forest types and plant lineages were treated as dependent binary variables, while the CWM of each plant trait was treated as a dependent continuous (Gaussian) variable for the models. For this, we trained repeated generalized linear models (GLMs) following a similar machine learning framework for modeling β-diversity. Overall, we first randomly split 60% of all forest inventories available per county for training purposes, following the same approach described above to ensure a balanced representation of forest inventories across counties. After data partitioning, we implemented a set of model configurations to evaluate the relative and combined contribution of predictors, including MDS β-diversity axes alone, climatic and topographic variables alone, and a combination of both. This approach allowed us to assess whether MDS axes capture sufficient ecological information to predict forest attributes independently, or whether additional environmental context improves model performance. The combined models were designed to test the complementarity between forest community composition and abiotic drivers, as MDS axes may implicitly encode environmental variation while also capturing biotic structure not fully explained by climate and topography alone.

For each model configuration, we train GLMs (n = 100) using 10-fold cross-validation for each forest type, plant lineage or plant trait using 85% of the training samples on each iteration. To overcome the imbalance of observations of scarce forest types and plant lineages, we applied the random over-sampling examples technique (Lunardon et al., 2014) on each iteration to aid the binary classification. Once the models were trained, we assessed their performance using I) training and II) testing datasets and using III) predicted MDS axes from spaceborne spectroscopy

on datasets that were used to train previous PLSR models, and IV) predicted MDS axes from spaceborne spectroscopy on datasets that were used validate previous PLSR models. For I and II, categorical models of forest type and plant lineages were evaluated using True Skill Statistic (TSS), sensitivity, specificity, and the Area Under the Curve (AUC) of the Receiver Operating Characteristic curve (ROC), while continuous models of CWM of plant traits were evaluated using R^2, bias, RMSE, and %RMSE. For III and IV, all models were evaluated using R^2, bias, RMSE, and %RMSE through the observed-predicted occurrence probability of forest type or plant lineages as well as the observed-predicted CWM of plant traits.

To quantify the contribution of individual predictors, we estimated the relative importance of variables for the full models (i.e., models including both MDS axes and climatic and topographic information). For each fitted GLM, relative importance was estimated using the 'varImp' function in caret (Kuhn, 2008), based on the absolute magnitude of model coefficients, and normalized to sum to one within each model. Relative importance values were then averaged across the 100 model iterations to obtain robust estimates for each predictor and forest attribute. Predictors were grouped into three categories—MDS axes, climatic variables, and topographic variables—and their contributions were aggregated to assess their overall importance. Finally, we quantified their climatic and topographic hypervolume to characterize the environmental breadth of forest types and plant lineages. Specifically, we first reduced the dimensionality of the climatic and topographic variables using a PCA, retaining the first three principal components to represent the dominant gradients of environmental variation. For each forest type or plant lineage, we then calculated the convex hull volume encompassing all occurrences within this reduced environmental space. This hypervolume was used as a proxy for the environmental niche breadth of each group and was subsequently related to the relative importance of predictor groups in the corresponding models.

2.9 Mapping β-diversity, forest attributes, and change.

We applied our predictive models to map dimensions of forest community composition using scenes of DESIS and EMIT (i.e., Level 3 products) and then map forest attributes associated with the probability of forest types, lineages and CWM of plant traits (i.e., Level 4 products). We first applied the repeated PLSR models on DESIS and EMIT scenes to generate maps of the mean estimated MDS axes associated with dimensions of community composition. The uncertainty of these estimates per MDS β-diversity axis and dimension of diversity were mapped as the amplitude between the upper and lower limits of the confidence intervals at 95% of the predicted values following (Guzmán Q. et al., 2023), as well as the sum of amplitudes among all axes from each dimension as a descriptor of the overall uncertainty. We then applied the coefficients from the repeated GLMs on these MDS maps in addition to climatic and topographic layers to generate maps of forest attributes and their uncertainties as described above.

We used EMIT scenes to showcase the application of our methods. We showcase the mapping of MDS β-diversity axis using a group of long-track scenes from 2025 from Southeast Ohio to Southeast Virginia. These scenes encompass documented transitions of forest community composition such as those from the FIA ecoregions (USDA Forest Service, 2023). For the study case of mapping of forest attributes, we showcase scenes from the southern Appalachian Mountains in Tennessee and North Carolina. These scenes encompass sections of one of the most biodiverse areas of North America, including the Great Smoky Mountains National Park, Nantahala National Forest, and Frozen Head State Park. We highlight probability predictions of three of the most abundant forest types and plant lineages according to FIA plots in the region as well as predictions of CWM of leaf mass per area, wood density, and leaf nitrogen concentration. Finaly, to evaluate the ability of predicted MDS β-diversity axes to capture temporal changes in forest community composition, we analyzed repeated EMIT imaging spectroscopy observations

acquired over a region in Arkansas in 2024 and 2025. To quantify changes in community composition, we compared the predicted MDS Pβ-diversity axes between years in both geographic and ordination space. In geographic space, pixel-wise differences in axis values were computed to identify areas of compositional change. In ordination space, pixels from each year were projected onto the same MDS axes, allowing direct comparison of their trajectories. Directional shifts in ordination space were assessed by examining the displacement of pixel clusters between years relative to the overall distribution of observations. Areas exhibiting notable deviations were further analyzed in relation to known land-use changes. The application of our mapping efforts of β-diversity and forest attributes to all the scenes used from DESIS and EMIT are available at Harvard Dataverse.

3. **Results**

3.1 Correspondence between β-diversity and spectral dissimilarity

We first evaluated the relationships between spectral dissimilarity—derived using Spectral Angle Mapper (SAM)—and dimensions of β_{total}, revealing that communities differing in composition also tend to exhibit dissimilar spectral signatures (Fig. 2). These associations were generally larger in spaceborne observations from DESIS (r = 0.15–0.25) (Fig. 2A, B, and C) than from EMIT (r = 0.05–0.06) (Fig. 2G, H, and I), as indicated by Mantel tests. Spectral dissimilarity among communities appeared to be greater when estimated from the full spectral range of EMIT observations than from DESIS, but it was similar between DESIS and EMIT when the comparison was restricted to the same spectral range. Within the DESIS dataset, correlations with spectral dissimilarity were slightly stronger for $P\beta_{total}$ (r = 0.25) and $F\beta_{total}$ (r = 0.22) than $T\beta_{total}$ (r = 0.15). Furthermore, the correspondence between spectral dissimilarity and forest composition was more pronounced when using β_{total} rather than their partitions (i.e., $\beta_{replacement}$ or $\beta_{richness}$) for both DESIS (Fig. S4) and EMIT (Fig. S5). Given the strong and consistent correlation of the dimensions of β_{total} with spectral dissimilarity, matrices of the former were used in the subsequent analyses.

3.2 β-diversity ordination patterns across the Eastern United States

To address the high dimensionality of the β_{total} matrices for modeling purposes, we applied Multi-Dimensional Scaling (MDS) as a data reduction technique, reducing β_{total} matrices of each dimension into three axes each. This approach revealed patterns of dissimilarity in forest community composition and their associated spatial distributions (Fig. 3). Although all ordinations were derived from the same inventory data, MDS axes based on $P\beta_{total}$ and $F\beta_{total}$ information exhibited stronger goodness-of-fit (R^2 = 0.71 and 0.60, respectively) than MDS axes based on $T\beta_{total}$ (R^2 = 0.11), when correlating pairwise ordination distances with their respective β_{total} matrices (Fig. S6). Nonetheless, MDS axes were highly correlated across diversity dimensions within the same ordination space (Fig. S7).

The projection of weighted average scores and the correlation of climatic and topographic variables with the MDS ordinations (Fig. S8–S9) revealed the presence of underlying climatic and topographic gradients that likely influence forest community compositional dissimilarity. Our analyses suggested that across all three ordinations, higher mean annual temperature (MAT) and lower values of temperature annual range (TAR) were associated with the higher values of the first MDS axes. In contrast, lower precipitation seasonality (PS) and higher mean annual precipitation (MAP) were associated with the second axes, while elevation and slope correlated with the third axes. The positive associations of the first two MDS axes with MAT and MAP are consistent with expectations based on the MDS rotation procedure (see Materials and Methods for details). Furthermore, the MDS ordinations also appeared to capture gradients in the CWM of

plant traits and the relative abundances of plant lineages, as indicated by the projection of weighted average scores and correlations with MDS axes (Fig. S10–S11).

3.3 Predicting ordinations of forest community composition from spectra

We developed a workflow (Fig. S12) to predict MDS ordination axes for each dimension of β_{total} using observations from DESIS and EMIT and by applying repeated Partial Least Squares Regression (PLSR) within a machine learning framework. Our results demonstrate that it is feasible to predict MDS β-diversity axes across all ordinations (Fig. S13 – S16), highlighting the potential to map multiple dimensions of forest community composition at large spatial scales. Model performance varied considerably across diversity dimensions, MDS β-diversity axes, and sensors, with training and testing R^2 values ranging from 0.08 to 0.64 and percent RMSE (%RMSE) between 17% and 29% (Tables S2 and S3). Across all models, Axes 1 consistently yielded higher R^2 than Axes 2 but not Axes 3. Predictions for Axes 1 require more latent components to achieve optimal predictions than Axes 2 or 3 (Fig. S17). Between sensors, models based on EMIT data often outperformed those based on DESIS (Table S2 and S3).

Despite the observed differences in model performance, the Variable Importance in Projection (VIP) from PLSR models revealed consistent spectral predictors across sensors, diversity dimensions, and MDS axes (Fig. 4). Spectral regions around 710, 790, and 1521 nm consistently contributed to the prediction of community composition across taxonomic, phylogenetic, and functional dimensions. These wavelengths are generally linked with vegetation greenness, forest structure, water content, respectively revealing potential spectral features that underpin key biophysical and biochemical aspects of community composition (Table S4).

We applied the PLSR coefficients at the optimal latent components (Fig. S18) to map dimensions of forest community composition across the Eastern U.S. at moderately high spatial resolution (30 m or 60 m). Here we illustrate their application on a group of long-track scenes from EMIT from Southeast Ohio to Southeast Virginia (Fig. 5); an area close to 43,830 km^2. This study region spans pronounced topographic variation (Fig. 5A) and corresponding changes in canopy reflectance captured by EMIT imagery (Fig. 5B). Predicted patterns of taxonomic, phylogenetic, and functional composition derived from MDS β-diversity ordinations (Fig. 5D, G, and J) closely mirror field-observed community gradients (Fig. 5C, F, and I). Map values and color scales correspond to those in Fig. 3, enabling the back-projection of mapped communities within the ordination space. Variation in forest community composition across this region—captured by the predicted MDS β-diversity axes—aligns with a transition from Midwestern assemblages to lowland coastal forests, with a clear turnover across the Appalachian Mountains. We then computed the uncertainty of prediction for each pixel in these composition maps (Fig. 5E, H, and K) as the summed amplitude of predictions, highlighting areas of lower model confidence. Among scenes, the uncertainty in the composition tends to co-vary across dimensions, where higher values are observed in lowland areas likely due to spectral mixing of frequent non-forest covers (e.g., Fig. S19). Overall, these results demonstrate how integrating spaceborne spectral data with β-diversity ordinations enables the mapping of dimensions of biodiversity, providing a scalable framework for monitoring forest composition and its environmental controls across large spatial extents. The following sections showcase how these maps could be used to predict forest attributes and to evaluate temporal changes in forest community composition.

3.4 Predicting and mapping forest attributes using β-diversity axes

We showcase one potential application of MDS β-diversity axes and their mapping to enhance the prediction of forest attributes in the context of distribution models. To this end, we evaluated whether MDS axes, with and without the inclusion of climatic and topographic variables, can

effectively predict forest attributes, including the presence or absence FIA-classified forest type or plant lineages, and CWM of plant traits. Specifically, the probability of occurrence of specific FIA-forest types was predicted using taxonomic ordinations and composition maps (Fig. S20), while ordinations and maps of phylogenetic composition informed predictions of plant lineage occurrence (Fig. S21). Similarly, ordinations and maps of functional composition enabled the estimation of CWM of plant traits (Fig. S22). Using generalized linear models (GLMs), we found that integrating the MDS axes with climatic and topographic variables consistently improved predictive performance relative to models based on axes alone or on climatic and topography alone (Fig. S23–S25). When evaluating 59 forest types, binomial models combining all predictors achieved an average True Skill Statistic (TSS) of 0.68 ± 0.14 and an Area Under the Curve (AUC) of 0.90 ± 0.06, outperforming models based solely on climatic and topographic variables (TSS = 0.57 ± 0.18; AUC = 0.84 ± 0.09) or on MDS axes alone (TSS = 0.59 ± 0.15; AUC = 0.85 ± 0.08). Likewise, when predicting the presence–absence of 28 plant lineages, models including all predictors yielded an average TSS of 0.54 ± 0.19 and AUC of 0.84 ± 0.10, exceeding the performance of models based only on climatic and topographic variables (TSS = 0.42 ± 0.20; AUC = 0.75 ± 0.10) or MDS axes alone (TSS = 0.40 ± 0.17; AUC = 0.75 ± 0.10). Finally, continuous models predicting CWMs of eight plant traits achieved an average R^2 of 0.46 ± 0.20 (%RMSE = 15.18 ± 2.03) when all predictors were included, outperforming models based solely on climatic and topographic variables (R^2 = 0.23 ± 0.15; %RMSE = 18.59 ± 1.37) or on MDS axes alone (R^2 = 0.36 ± 0.21; %RMSE = 16.53 ± 2.12).

Given that models integrating MDS axes with climatic and topographic variables on average yielded the highest performance across forest attributes, we focused subsequent analyses on these full models' configuration to assess the relative importance of individual predictors. Overall, the relative importance of variables in these GLMs indicated that three MDS axes derived from β_{total} dimensions accounted for more than 48% (±13.11%) of the total variable importance across forest attributes. In contrast, four climatic and three topographic variables contributed 36.48% (±11.35%) and 15.11% (±7.96%), respectively (Fig. S26–S28). The relative importance of these three MDS axes appeared to depend on the forest attribute in question. On average, the forest attribute was more relevant in CWM plant trait models (55.72 ± 5.13 %) than in presence or absence models of plant lineages (50.14 ± 16.70 %) or in forest type models (46.59 ± 11.37 %). In these latter models, we observed that the MDS axes had greater relative importance in predicting forest types and plant lineages with high climatic and topographic hypervolume—defined as the convex hull volume of their occurrences within the reduced multivariate space of climate and topography represented by the first three principal components. Consequently, climatic and topographic variables had lower relative importance in predicting forest types and plant lineages with low climatic and topographic hypervolumes (Fig. S29). In other words, forest types and plant lineages that span large climatic and topographic variation are better predicted when MDS axes are included.

By applying GLM coefficients to MDS β-diversity axes predicted from spaceborne spectroscopy at the overlapping inventory locations, we demonstrate that our workflow can infer a wide array of forest attributes from satellite data (Fig. 6; Fig. S30–S32). We evaluated the accuracy of these predictions by regressing expected probabilities (for categorical attributes) or CWMs (for continuous traits) from the GLM models against their predicted values. Using this validation approach across all testing datasets and both sensor platforms, an average R^2 of 0.69 ± 0.23 (%RMSE = 16.98 ± 5.51) was obtained for forest types, 0.70 ± 0.22 (%RMSE = 15.17 ± 7.58) for plant lineages, and 0.49 ± 0.24 (%RMSE = 16.83 ± 3.16) for community plant traits.

Although sensor-specific performance varied for some forest attributes, overall model accuracy was relatively consistent between sensors.

3.5 Case study I: predicting forest attributes across Southern Appalachia

Here we illustrate the application of our framework for mapping forest attributes using EMIT L2A scenes from the southern Appalachian Mountains, spanning parts of North Carolina and Tennessee, U.S. (Fig. 7). In this region, variation in forest community composition—captured by predicted MDS β-diversity axes—aligns closely with elevation gradients. Uncertainty of predictions appeared to be associated with lowland areas, likely due to the landscape heterogeneity and their role in spectral mixing with non-forest covers (e.g., Fig. S19). Leveraging the maps of community composition in combination with climatic and topographic variables, we then generated spatial predictions of various forest attributes. In this region, the high abundance of broadleaf deciduous communities at lower elevations (i.e., below ~1,000 m a.s.l.) corresponds with the high predicted occurrence of *Quercus* (oak) species and oak-dominated forest types in the scene. These areas are also characterized by elevated foliar nitrogen (N) concentrations and lower predicted values of leaf mass per area (LMA). In contrast, at higher elevations—primarily above 1,800 m a.s.l.—there is a greater prevalence of evergreen coniferous communities, marked by high predicted occurrence of *Picea* (spruce) species, high LMA values, and lower foliar N concentrations. Mid-elevational zones exhibit a transitional pattern, with forest communities dominated by *Acer saccharum* (sugar maple), *Betula alleghaniensis* (yellow birch), and *Fagus grandifolia* (American beech) to mention some.

3.6 Case study II: mapping temporal changes in forest community composition

We showcase a second potential application of MDS β-diversity axes maps to detect spatiotemporal changes in forest community composition. Using repeated EMIT observations acquired in 2024 and 2025 over a region in Arkansas, we generated maps of phylogenetic community composition for each year and quantified their differences in both geographic and ordination space (Fig. 8). The predicted MDS axes revealed coherent spatial patterns of community composition that were largely consistent between years (Fig. 8B and E). However, localized deviations were evident (Fig. 8H), particularly within the delineated polygons that correspond to evergreen plantations that were partially or totally harvested in 2025, where shifts of values in the mapped axes indicated measurable changes in community composition. These changes were further supported by the back-projection of pixels in ordination space (Fig. 8C, F, and I), where subsets of pixels associated with these polygons exhibited clear directional shifts between 2024 and 2025 relative to the broader regional observations. The distribution of pixels in ordination space (Fig. 8C and F) for 2024, interpreted considering the weighted average projections (i.e., Fig. S8), may suggest that pixels within the highlighted areas are associated with communities exhibiting greater affinity to gymnosperms, higher leaf mass per area, and warmer environmental conditions. These results showcase that predicted maps of β-diversity axes provide a consistent and interpretable representation of forest community composition that can be compared across time to identify areas of compositional change.

4. Discussion

We have developed a comprehensive framework for mapping dimensions of forest diversity by integrating thousands of forest inventory plots with spaceborne imaging spectroscopy. Our findings demonstrate that across the Eastern U.S., forest communities with dissimilar compositions (e.g., high β-diversity) exhibit corresponding spectral dissimilarities. This correspondence is consistent across dimensions of β-diversity (i.e., taxonomic, phylogenetic, and functional) and spaceborne sensors (i.e., DESIS and EMIT), providing a solid foundation for

using *in-situ* β-diversity to remotely sense forest diversity from space. Due to their high dimensionality, however, β-diversity matrices are challenging to apply directly to modeling biodiversity from satellite spectra. To address this challenge, we used ordinations of β-diversity, in which the combination of three axes captures spatial patterns of forest community composition. We show that by linking ordination axes of β-diversity with spaceborne spectroscopy, it is feasible to map multiple dimensions of forest community composition at large spatial scales. We further demonstrate that by integrating a range of forest attributes with maps of forest community composition, our framework can be used to map the distribution of forest types, plant lineages, and community plant traits. This approach reveals that forest community composition—captured through β-diversity axes—contributes more strongly than climate or topography to predicting different forest attributes, likely because remotely sensed and inventory-based data capture local variation that is not resolved by broader-scale environmental variables (Pinto-Ledezma and Cavender-Bares, 2020). In this sense, climate and topography set the biogeographic context of which forest attributes may occur, whereas spectroscopy data capture what is present on the ground, offering a more ecologically grounded basis for mapping species occurrence and functional traits. We also demonstrate that maps of forest community composition can be used to evaluate temporal changes in forest cover, highlighting their potential for monitoring forest diversity. The effectiveness of our framework for mapping forest community composition and diverse forest attributes at high spatial resolution and across large spatial extents sets the stage for verification of stewardship efforts aimed at sustaining planetary biodiversity and resilience. The framework components are discussed in detail below.

4.1 Correspondence between β-diversity and spectral dissimilarity.

The consistent correspondence between spectral dissimilarity and β-diversity across spaceborne sensors and diversity dimensions supports our central hypothesis that forest communities differing in taxonomic, phylogenetic, or functional composition exhibit greater spectral dissimilarity than communities with similar compositions. These community-level findings are congruent with previous species-level (Griffith et al., 2023; Schweiger et al., 2018) or community-level (Everest et al., 2025; Gholizadeh et al., 2020; Schweiger and Laliberté, 2022) studies using airborne imaging spectroscopy that show closely related or functionally similar species or communities have similar reflectance spectra. Our findings further indicate that differences in spatial and spectral range among sensors play an important role in the nature of the relationship between vegetation and spectral similarity. For example, the greater spectral dissimilarity among communities detected by EMIT may be attributable to its broader spectral range and coarse spatial resolution compared with DESIS. When forest communities exhibit limited spectral mixing with other cover types (e.g., grassland, shrubs, agriculture fields), this broader spectral range allows the detection of a wider suite of biochemical and structural canopy properties that are not fully captured in the visible–near infrared region (Wang et al., 2020), increasing the magnitude of spectral differentiation among contrasting forest communities. Conversely, DESIS provides higher correspondence when finer spatial resolution is critical, as indicated by comparisons of DESIS and EMIT using the same spectral region (i.e., 30 versus 60 m). Finer spatial resolution has the potential to enhance spectral sensitivity to local compositional turnover by reducing pixel-level mixing among co-occurring species and other non-forest land cover, strengthening the relationship between spectral dissimilarity and β-diversity.

Moreover, our findings reveal that the strength of the correspondence between β-diversity and spectral dissimilarity depends not only on sensor properties but also on how β-diversity and their partitions are estimated. The spectral correspondence across dimensions and partitions of β-

diversity suggests that the combined effects of $\beta_{replacement}$ and $\beta_{richness}$—captured by β_{total}—are more consistently and strongly associated with the spectral dissimilarity among communities than either partition alone. Nevertheless, spectral dissimilarity appears to capture shared patterns across taxonomic, phylogenetic, and functional dimensions of β_{total} without favoring any dimension. These results indicate that β_{total} integrates multiple sources of compositional variation that jointly influence the optical properties of communities, making it particularly well suited for remote sensing applications. Collectively, our results highlight the importance of sensor properties and diversity metrics in shaping observed relationships between β-diversity and spectral dissimilarity, reinforcing the potential of using β_{total} as a congruent approach for quantifying forest diversity through remotely sensed measures of community dissimilarity.

4.2 β-diversity patterns across the Eastern United States

Our framework uses ordinations of β_{total} to summarize the complexity of large community dissimilarity matrices and to reveal spatial patterns of community dissimilarity across broad spatial extents. The β-diversity patterns captured by these ordinations broadly recapitulate the spatial clustering of forest community composition represented by FIA eco-provinces (McNab et al., 2007; Ruefenacht et al., 2008). However, across the Eastern U.S., these ordinations describe the spatial variation in forest community composition within a continuous ordination space that emphasizes gradients of community differentiation rather than discrete regional boundaries. Visually, this continuous ordination space is particularly useful for differentiating communities with contrasting compositions that are closely located in geographic space (e.g., in the Southeastern U.S.). The strong correlation among MDS axes across dimensions further indicates that the relative positioning of communities in the ordination space is largely conserved, implying that MDS axes capture congruent patterns of community differentiation. Such correlations also imply that taxonomic, phylogenetic, and functional β_{total} are tightly coupled, likely reflecting common underlying environmental and biogeographic processes shaping forest community composition as suggested by other studies (Pinto-Ledezma et al., 2018; Qian et al., 2013).

Despite the strong correlation of MDS axes among dimensions, the higher goodness-of-fit observed in phylogenetic and functional β_{total} ordinations suggests that the compositional dissimilarity of communities is more effectively represented when evolutionary and functional relationships among communities are considered, rather than relying solely on taxonomic composition. In other words, phylogenetic and functional dimensions may integrate shared evolutionary constraints or ecological strategies that are more informative in ordinating communities than species identities alone. These findings support the interpretation that MDS ordinations of β-diversity provide an effective framework for characterizing dimensions of forest community composition across broad spatial extents, with the potential to be linked to remote sensing observations for mapping forest diversity.

4.3 Predicting ordinations of forest community composition from spectra

Our modeling framework demonstrates that MDS ordination axes derived from β_{total} across taxonomic, phylogenetic, and functional dimensions can be predicted from spaceborne imaging spectroscopy, supporting the feasibility of mapping multiple facets of forest community composition at large spatial scales. Although predictive performance varied across diversity dimensions, ordination axes, and sensors, the consistent ability to recover signal in the dominant axes—particularly Axis 1—suggests that major gradients of community differentiation are spectrally expressed. Lower and more variable performance for Axes 2 and 3 likely reflects their representation of subtler or more localized compositional gradients, which may exhibit a weaker

association with the optical properties of communities or be confounded by noise introduced through spatial mixing, environmental heterogeneity, and uncertainty in both inventory and spectral data. Together, these results highlight that while spaceborne spectroscopy is effective at capturing broad-scale patterns of community composition, predicting finer-scale or higher-order axes of MDS β-diversity remains challenging and may require higher resolution imagery, additional structural information such as LiDAR, or complementary data sources.

Despite these differences in predictive performance, the strong consistency of Variable Importance in Projection (VIP) scores across sensors, diversity dimensions, and MDS axes underscores the robustness of the spectral–diversity relationships captured by our models. The alignment of VIP profiles between DESIS and EMIT indicates that signals of forest community composition are not sensor-specific but instead are associated with key spectral regions that are broadly detectable from space. The recurrent importance of wavelengths linked to vegetation greenness, canopy structure, and water content suggests that these biophysical and biochemical properties form a common spectral basis underlying multiple dimensions of β-diversity. Moreover, the similarity of VIP patterns across ordination axes implies that ecologically meaningful gradients of community composition are consistently expressed in the spectral space, even when their predictability varies. These results point toward the potential for transferring models across sensors and the development of sensor-agnostic approaches for large-scale biodiversity monitoring, which will likely strengthen the feasibility of operational forest diversity mapping using current and forthcoming spaceborne spectroscopy missions such as NASA's SBG (Surface Biology and Geology), NASA's EAGLE-VSWIR, and ESA's CHIME (Copernicus Hyperspectral Imaging Mission for the Environment).

4.4 Case study I: predicting forest attributes

We showcase the potential of using MDS β-diversity axes to predict forest attributes and to map their occurrence probability (e.g., forest types or plant lineages) or their community traits within a distribution modeling framework. Among model comparisons, the integration of MDS β-diversity axes with climate and topographic predictors consistently yielded the highest performance, outperforming models based solely on climatic and topographic variables or on MDS β-diversity axes alone. Furthermore, greater explanatory power of MDS β-diversity axes when compared with climatic and topographic variables underscores the importance of community context for predicting a broad range of forest attributes. In essence, compositional information captured by ordinations provides complementary ecological information beyond that explained by abiotic predictors alone. By capturing multidimensional β-diversity among communities, ordinations can reflect not only community gradients but also the outcomes of species interactions and ecological filtering embedded within assemblages. In this sense, forest attributes derived from forest community composition, climatic, and topographic information reflect the combined influence of realized species distributions, assemblage structure, and abiotic constraints. Within a distribution modeling context, therefore, interpreting forest community composition through spaceborne spectroscopy provides a means to capture information inherent in realized assemblages, rather than strict representations of fundamental niches as is commonly done (Hutchinson, 1978; Silvertown, 2004).

Although MDS β-diversity axes capture substantial ecological information, climatic and topographic variables remain essential for representing the broader environmental constraints shaping the distribution of forest attributes. This is particularly important for some forest types and plant lineages with lower climatic and topographic hypervolumes, suggesting that their distributions are restricted to narrower environmental conditions. Under such conditions, abiotic variables are likely to exert stronger control over their occurrence, thereby prevailing over MDS β-diversity axes in explaining their distribution. Incorporating these abiotic predictors therefore

helps anchor community composition within their underlying environmental context, improving the discrimination of forest attributes that are tightly associated with specific climatic or topographic settings. In our framework, climate and topography appear to provide macroecological boundaries in which community assembly unfolds, while MDS β-diversity ordinations capture the realized configuration of assemblages within those boundaries. In general, these results showcase the potential use of MDS β-diversity ordinations predicted from spaceborne observations to generate ecologically meaningful, spatially continuous predictions of forest community composition and their attributes—paving the way for scalable biodiversity assessments from space.

4.5 Case study II: mapping temporal changes in forest community composition

The observed shifts in MDS phylogenetic β-diversity axes between 2024 and 2025 showcase the potential of this framework to capture meaningful changes in forest community composition. Because these axes summarize multidimensional patterns of community dissimilarity, their temporal displacement reflects coordinated changes in community assemblages, such as those induced by harvesting, regrowth, or shifts in species dominance. Between observations, shifts in mapped axis values do not necessarily correspond to changes in individual species per se, but rather to changes in the relative composition of communities within the ordination space. In the Arkansas case study, the directional shifts associated with harvested evergreen plantations are consistent with transitions away from gymnosperm-dominated assemblages, illustrating how ordination space can provide a mechanistic interpretation of changes in community composition. The interpretation of these changes should therefore be grounded in the ecological interpretation of MDS axes as representations of multidimensional β-diversity. In this context, the evaluation of spatiotemporal changes in forest community composition is particularly well suited to applications in ecosystem monitoring and disturbance assessment, where detecting shifts in composition is critical for understanding resilience, succession, or degradation processes.

In contrast to traditional approaches that depend on multiple reflectance bands and vegetation indices to assess change, MDS axes offer the advantage of dimensionality reduction, facilitating comparisons and enabling the detection of change by integrating information across the full spectral domain into a reduced set of ecologically interpretable gradients. However, several limitations should be acknowledged. The interpretation of MDS axes depends on their calibration and ecological grounding, which may vary across regions or datasets, potentially limiting transferability and broader application. Additionally, factors affecting spectral signals (e.g., spectral mixing and atmospheric conditions) can introduce uncertainty or confound compositional inference. Consequently, MDS-based mapping should be viewed as a complementary approach, most powerful when integrated with field data to support the contextualization of compositional changes.

4.6 Mapping forest diversity using spaceborne spectroscopy

The maps generated through our framework demonstrate that spaceborne imaging spectroscopy can provide spatially explicit, ecologically grounded information on multiple dimensions of forest community composition. In the context of increasing global commitments to biodiversity monitoring to verify efforts towards the UN Kunming–Montreal Global Biodiversity Framework goals and targets, there is a growing need for coordinated observation systems capable of integrating diverse data streams into consistent, scalable assessments of ecological change (Zurell et al., 2025). Our approach contributes to advancing these capabilities by enabling the characterization of forest diversity at the community level across broad spatial extents, translating complex β-diversity patterns into interpretable axes that capture taxonomic, phylogenetic, and

functional dimensions of biodiversity. The dimensions of forest diversity can be directly linked to conservation-relevant attributes, including the identification of forest types, the distribution of evolutionary lineages, the spatial organization of community functional traits, or even efforts associated with observing changes in community composition. The computational framework we have developed thus has the potential to support stewardship goals at politically relevant scales (Gonzalez et al., 2023). By leveraging spaceborne observations, our approach complements—rather than replaces—traditional field inventories, offering a scalable means for remotely detecting dimensions of forest diversity and their changes over time.

Although we demonstrate the feasibility of using spaceborne spectroscopy to map forest diversity, its practical application must be guided by the inherent limitations of both the technology and the framework. For instance, the spatial resolution of current spaceborne imaging spectroscopy (30–60 m) inherently leads to spectral mixing, whereby individual pixels represent integrated signals from multiple species and, in some cases, non-forest elements. While this limitation is consistent with the scale of many forest inventory plots and can be partially mitigated through filtering and uncertainty assessment, such mitigation requires constraining the interpretation of predictions according to user needs. Similarly, our analysis captures forest composition during a defined phenological window rather than resolving temporal dynamics, reflecting both the limited revisit frequency of current hyperspectral missions and the asynchronous nature of ground-based inventories. Thus, the integration of temporal dynamics in canopy observations for detecting β-diversity should be further evaluated, as it may influence the results (Everest et al., 2025). Additional complexities may arise from confounding factors such as disturbance, land-use history, and biotic agents that may influence both spectral signals and predictions of community composition as well as potentially bias model evaluation if not properly addressed. Although these confounding factors are implicitly represented in the observed data, their explicit attribution remains beyond the scope of the present framework and highlights important directions for future research.

The framework we have developed enables scalable and integrative forest diversity mapping as part of an emerging, globally distributed, multi-sensor biodiversity observation system (Navarro et al., 2017). While spaceborne spectroscopy alone cannot capture all dimensions of forest diversity—such as fine-scale structural variation or intraspecific trait diversity—it provides a unique and powerful lens on community-level composition that can be combined with complementary data sources, including LiDAR, multispectral time series, and expanded ground inventories. Our results demonstrate that forest community composition, captured by β-diversity ordinations, offers a robust and ecologically meaningful basis for observing land use change or predicting forest attributes, often exceeding the explanatory power of climate and topography alone. In general, the integration of spectral data, β-diversity ordinations, and environmental predictors opens new opportunities for generating spatially continuous, interpretable maps of biodiversity. Such advances have the potential to support conservation planning, guide field-based efforts, and ultimately contribute to the monitoring and verification of biodiversity outcomes at regional to global scales.

5. Conclusion

Our study presents a comprehensive methodological framework for mapping dimensions of forest diversity by integrating thousands of forest inventory plots with spaceborne imaging spectroscopy. Spaceborne spectroscopy offers considerable potential for enhancing our ability to observe and monitor forest biodiversity. Yet mapping forest diversity across large spatial extents—given diverse ecosystem patterns, high species richness, and major shifts in species composition—presents significant challenges. Our findings demonstrate that across the Eastern U.S., forest communities with dissimilar compositions (e.g., high β-diversity) exhibit corresponding spectral dissimilarities. The correspondence between spectral dissimilarity and

compositional dissimilarity is consistent across multiple dimensions of β-diversity (i.e., taxonomic, phylogenetic, and functional), for both components of β-diversity partitioning (i.e., replacement and richness), and for spectral information from two spaceborne sensors (i.e., DESIS and EMIT), providing a solid foundation for using *in situ* β-diversity to remotely sense forest diversity from space. Due to their complexity and high dimensionality, however, pairwise β-diversity matrices are challenging to apply directly to modeling biodiversity from satellite spectra. To address this challenge, our framework employs ordinations of β-diversity, in which the combination of three axes captures spatial patterns of forest community composition. We show that ordination axes of β-diversity can readily be used to predict biodiversity from imaging spectroscopy, and that their application to satellite observations allows the spatial mapping of multiple dimensions of forest community composition from space. Moreover, our modeling approach reveals regions of the electromagnetic spectrum that are consistently important for predicting axes of β-diversity across multiple diversity dimensions and for both spaceborne sensors, underscore the potential for models based on imaging spectroscopy to capture signals of forest composition and to generate a sensor-agnostic framework for large-scale diversity mapping. By integrating a range of forest attributes with axes of β-diversity across different diversity dimensions, along with climate and topographic variables, we showcase that our framework can effectively map forest types, plant lineages, and community plant traits. This approach reveals that forest community composition—captured through β-diversity axes—plays a stronger role than climate or topography, thus offering a more ecologically grounded basis for mapping species occurrence and functional traits. We also showcase how our framework could be used to evaluate temporal changes in forest community composition. The effectiveness of our framework for mapping forest community composition and diverse forest attributes at high spatial resolution and across large spatial extents sets the stage for potential verification of forest management efforts aimed at sustaining planetary biodiversity and resilience.

CRediT authorship contribution statement

J. Antonio Guzmán Q: Conceptualization, Data curation, Formal analysis, Software, Investigation, Methodology, Visualization, Validation, Writing – original draft, Writing—review & editing. **Jonathan A. Knott:** Data curation, Resources, Writing—review & editing. **Jesús N. Pinto-Ledezma:** Investigation, Writing – review & editing. **Philip A. Townsend:** Writing – review & editing. **Jeannine Cavender-Bares:** Conceptualization, Writing – original draft, Writing – review & editing, Resources, Supervision, Funding acquisition, Project administration.

Declaration of Competing Interest

Authors declare that they have no competing interests.

Acknowledgments

We thank the German Aerospace Center (DLR) and the National Aeronautics and Space Administration (NASA) for providing access to DESIS and EMIT imagery for scientific research, respectively. We thank the numerous field crews and information management staff who collected and compiled the FIA data used in this study. We thank Vinicius Marcilio-Silva and Steven Augustine for insightful discussions on imputing plant traits and ecological communities. Thanks to the Minnesota Supercomputing Institute and the Faculty of Arts and Sciences Research Computing for their high-performance infrastructure that allows us to conduct this research. ChatGPT was used to improve the English of the manuscript and debug the code. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. This study was supported by the NASA Biodiversity Program (Award number: 80NSSC21K1349), the National Science Foundation (NSF) through ASCEND Biology

Integration Institute (DBI: 2021898), the Harvard University Faculty of Arts and Sciences, and the Harvard University Salata Institute for Climate and Sustainability. J.N.P.-L. and P.A.T were supported by the NSF Division of Environmental Biology (DEB: 2017843) and NASA's EMIT Science program (Award number: 80NSSC24K0777), respectively.

Data availability

Forest inventory data were obtained from the USDA Forest Service's FIA Program and are available through the FIA DataMart (https://apps.fs.usda.gov/fia/datamart/datamart.html). However, as noted in the Methods, we used a federally protected version of the FIA database to access actual plot locations for our analyses (for more information on federally protected FIA data, see https://research.fs.usda.gov/programs/fia/sds). All code associated with this research is available on GitHub (https://github.com/Antguz/mapping-communities), and will be archived in Zenodo under version 1.0 upon publication. Data that do not compromise federally protected information are being prepared for release in the Harvard Dataverse. The spaceborne data products developed in this research are also available through the Harvard Dataverse (https://dataverse.harvard.edu/previewurl.xhtml?token=cfb44b92-ec7f-4cd8-9c7c-c2b4b43612d6).

References

Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19, 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x

Bechtold, W.A., Patterson, P.L., 2015. The Enhanced Forest Inventory and Analysis Program - National Sampling Design and Estimation Procedures (No. SRS-GTR-80). U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC. https://doi.org/10.2737/SRS-GTR-80

Bolton, D.K., Gray, J.M., Melaas, E.K., Moon, M., Eklundh, L., Friedl, M.A., 2020. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. Remote Sensing of Environment 240, 111685. https://doi.org/10.1016/j.rse.2020.111685

Brooks, B.-G.J., Lee, D.C., Pomara, L.Y., Hargrove, W.W., 2020. Monitoring Broadscale Vegetational Diversity and Change across North American Landscapes Using Land Surface Phenology. Forests 11, 606. https://doi.org/10.3390/f11060606

Cardoso, P., Rigal, F., Carvalho, J.C., 2015. BAT – Biodiversity Assessment Tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. Methods Ecol Evol 6, 232–236. https://doi.org/10.1111/2041-210X.12310

Cardoso, P., Rigal, F., Carvalho, J.C., Fortelius, M., Borges, P.A.V., Podani, J., Schmera, D., 2014. Partitioning taxon, phylogenetic and functional beta diversity into replacement and richness difference components. Journal of Biogeography 41, 749–761. https://doi.org/10.1111/jbi.12239

Cavender-Bares, J., Schneider, F.D., Santos, M.J., Armstrong, A., Carnaval, A., Dahlin, K.M., Fatoyinbo, L., Hurtt, G.C., Schimel, D., Townsend, P.A., Ustin, S.L., Wang, Z., Wilson, A.M., 2022. Integrating remote sensing with ecology and evolution to advance biodiversity conservation. Nat Ecol Evol 6, 506–519. https://doi.org/10.1038/s41559-022-01702-5

Cavender-Bares, J.M., Nelson, E., Meireles, J.E., Lasky, J.R., Miteva, D.A., Nowak, D.J., Pearse, W.D., Helmus, M.R., Zanne, A.E., Fagan, W.F., Mihiar, C., Muller, N.Z., Kraft, N.J.B., Polasky, S., 2022. The hidden value of trees: Quantifying the ecosystem services of tree lineages and their major threats across the contiguous US. PLOS Sustain Transform 1, e0000010. https://doi.org/10.1371/journal.pstr.0000010

Dalmayne, J., Möckel, T., Prentice, H.C., Schmid, B.C., Hall, K., 2013. Assessment of fine-scale plant species beta diversity using WorldView-2 satellite spectral dissimilarity. Ecological Informatics 18, 1–9. https://doi.org/10.1016/j.ecoinf.2013.05.004

Draper, F.C., Baraloto, C., Brodrick, P.G., Phillips, O.L., Martinez, R.V., Honorio Coronado, E.N., Baker, T.R., Zárate Gómez, R., Amasifuen Guerra, C.A., Flores, M., Garcia Villacorta, R., V. A. Fine, P., Freitas, L., Monteagudo-Mendoza, A., J. W Brienen, R., Asner, G.P., 2019. Imaging spectroscopy predicts variable distance decay across contrasting Amazonian tree communities. Journal of Ecology 107, 696–710. https://doi.org/10.1111/1365-2745.13067

Everest, J.J., Van Cleemput, E., Beamish, A.L., Spasojevic, M.J., Humphries, H.C., Elmendorf, S.C., 2025. Evaluating the utility of hyperspectral data to monitor local-scale β-diversity across space and time. Remote Sensing of Environment 316, 114507. https://doi.org/10.1016/j.rse.2024.114507

Féret, J.-B., Asner, G.P., 2014. Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. Ecological Applications 24, 1289–1296. https://doi.org/10.1890/13-1824.1

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Intl Journal of Climatology 37, 4302–4315. https://doi.org/10.1002/joc.5086

Gholizadeh, H., Gamon, J.A., Helzer, C.J., Cavender-Bares, J., 2020. Multi-temporal assessment of grassland α- and β-diversity using hyperspectral imaging. Ecological Applications 30, e02145. https://doi.org/10.1002/eap.2145

Gholizadeh, H., Gamon, J.A., Townsend, P.A., Zygielbaum, A.I., Helzer, C.J., Hmimina, G.Y., Yu, R., Moore, R.M., Schweiger, A.K., Cavender-Bares, J., 2019. Detecting prairie biodiversity with airborne remote sensing. Remote Sensing of Environment 221, 38–49. https://doi.org/10.1016/j.rse.2018.10.037

Gonzalez, A., Vihervaara, P., Balvanera, P., Bates, A.E., Bayraktarov, E., Bellingham, P.J., Bruder, A., Campbell, J., Catchen, M.D., Cavender-Bares, J., Chase, J., Coops, N., Costello, M.J., Czúcz, B., Delavaud, A., Dornelas, M., Dubois, G., Duffy, E.J., Eggermont, H., Fernandez, M., Fernandez, N., Ferrier, S., Geller, G.N., Gill, M., Gravel, D., Guerra, C.A., Guralnick, R., Harfoot, M., Hirsch, T., Hoban, S., Hughes, A.C., Hugo, W., Hunter, M.E., Isbell, F., Jetz, W., Juergens, N., Kissling, W.D., Krug, C.B., Kullberg, P., Le Bras, Y., Leung, B., Londoño-Murcia, M.C., Lord, J.-M., Loreau, M., Luers, A., Ma, K., MacDonald, A.J., Maes, J., McGeoch, M., Mihoub, J.B., Millette, K.L., Molnar, Z., Montes, E., Mori, A.S., Muller-Karger, F.E., Muraoka, H., Nakaoka, M., Navarro, L., Newbold, T., Niamir, A., Obura, D., O'Connor, M., Paganini, M., Pelletier, D., Pereira, H., Poisot, T., Pollock, L.J., Purvis, A., Radulovici, A., Rocchini, D., Roeoesli, C., Schaepman, M., Schaepman-Strub, G., Schmeller, D.S., Schmiedel, U., Schneider, F.D., Shakya, M.M., Skidmore, A., Skowno, A.L., Takeuchi, Y., Tuanmu, M.-N., Turak, E., Turner, W., Urban, M.C., Urbina-Cardona, N., Valbuena, R., Van De Putte, A., Van Havre, B., Wingate, V.R., Wright, E., Torrelio, C.Z., 2023. A global biodiversity observing system to unite monitoring and guide action. Nat Ecol Evol 7, 1947–1952. https://doi.org/10.1038/s41559-023-02171-0

Griffith, D.M., Byrd, K.B., Anderegg, L.D.L., Allan, E., Gatziolis, D., Roberts, D., Yacoub, R., Nemani, R.R., 2023. Capturing patterns of evolutionary relatedness with reflectance spectra to model and monitor biodiversity. Proc. Natl. Acad. Sci. U.S.A. 120, e2215533120. https://doi.org/10.1073/pnas.2215533120

Gu, H., Singh, A., Townsend, P.A., 2015. Detection of gradients of forest composition in an urban area using imaging spectroscopy. Remote Sensing of Environment 167, 168–180. https://doi.org/10.1016/j.rse.2015.06.010

Guzmán Q., J.A., Pinto-Ledezma, J.N., Frantz, D., Townsend, P.A., Juzwik, J., Cavender-Bares, J., 2023. Mapping oak wilt disease from space using land surface phenology. Remote Sensing of Environment 298, 113794. https://doi.org/10.1016/j.rse.2023.113794

Hargrove, W.W., Hoffman, F.M., 1999. Using multivariate clustering to characterize ecoregion borders. Comput. Sci. Eng. 1, 18–25. https://doi.org/10.1109/5992.774837

Hutchinson, G.E., 1978. An introduction to population ecology. Yale University Press, New Haven.

IPBES, 2019. Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Zenodo. https://doi.org/10.5281/ZENODO.3831673

Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F.W., Asner, G.P., Guralnick, R., Kattge, J., Latimer, A.M., Moorcroft, P., Schaepman, M.E., Schildhauer, M.P., Schneider, F.D., Schrodt, F., Stahl, U., Ustin, S.L., 2016. Monitoring plant functional diversity from space. Nature Plants 2, 16024. https://doi.org/10.1038/nplants.2016.24

Jin, Y., Qian, H., 2022. V.PhyloMaker2: An updated and enlarged R package that can generate very large phylogenies for vascular plants. Plant Diversity 44, 335–339. https://doi.org/10.1016/j.pld.2022.05.005

Knott, J.A., Liknes, G.C., Giebink, C.L., Oh, S., Domke, G.M., McRoberts, R.E., Quirino, V.F., Walters, B.F., 2023. Effects of outliers on remote sensing-assisted forest biomass estimation: A case study from the United States national forest inventory. Methods Ecol Evol 14, 1587–1602. https://doi.org/10.1111/2041-210X.14084

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. Remote Sensing of Environment 44, 145–163. https://doi.org/10.1016/0034-4257(93)90013-N

Krutz, D., Müller, R., Knodt, U., Günther, B., Walter, I., Sebastian, I., Säuberlich, T., Reulke, R., Carmona, E., Eckardt, A., Venus, H., Fischer, C., Zender, B., Arloth, S., Lieder, M., Neidhardt, M., Grote, U., Schrandt, F., Gelmi, S., Wojtkowiak, A., 2019. The Instrument Design of the DLR Earth Sensing Imaging Spectrometer (DESIS). Sensors 19, 1622. https://doi.org/10.3390/s19071622

Kuhn, M., 2008. Building Predictive Models in *R* Using the **caret** Package. J. Stat. Soft. 28. https://doi.org/10.18637/jss.v028.i05

Kuhn, M., Johnson, K., 2013. Applied predictive modeling, 1st ed. Springer, New York.

Leclère, D., Obersteiner, M., Barrett, M., Butchart, S.H.M., Chaudhary, A., De Palma, A., DeClerck, F.A.J., Di Marco, M., Doelman, J.C., Dürauer, M., Freeman, R., Harfoot, M., Hasegawa, T., Hellweg, S., Hilbers, J.P., Hill, S.L.L., Humpenöder, F., Jennings, N., Krisztin, T., Mace, G.M., Ohashi, H., Popp, A., Purvis, A., Schipper, A.M., Tabeau, A., Valin, H., Van Meijl, H., Van Zeist, W.-J., Visconti, P., Alkemade, R., Almond, R., Bunting, G., Burgess, N.D., Cornell, S.E., Di Fulvio, F., Ferrier, S., Fritz, S., Fujimori, S., Grooten, M., Harwood, T., Havlík, P., Herrero, M., Hoskins, A.J., Jung, M., Kram, T., Lotze-Campen, H., Matsui, T., Meyer, C., Nel, D., Newbold, T., Schmidt-Traub, G., Stehfest, E., Strassburg, B.B.N., Van Vuuren, D.P., Ware, C., Watson, J.E.M., Wu, W., Young, L., 2020. Bending the curve of terrestrial biodiversity needs an integrated strategy. Nature 585, 551–556. https://doi.org/10.1038/s41586-020-2705-y

Lunardon, N., Menardi, G., Torelli, N., 2014. ROSE: a Package for Binary Imbalanced Learning. The R Journal 6, 79. https://doi.org/10.32614/RJ-2014-008

McNab, W.H., Cleland, D.T., Freeouf, J.A., Keys, J.E., Nowacki, G.J., Carpenter, C.A., 2007. Description of "Ecological Subregions: Sections of the Conterminous United States" (No. Gen. Tech. Report WO-76B). Forest Service, United States Department of Agriculture.

McRoberts, R.E., Chen, Q., Walters, B.F., Kaisershot, D.J., 2018. The effects of global positioning system receiver accuracy on airborne laser scanning-assisted estimates of aboveground biomass. Remote Sensing of Environment 207, 42–49. https://doi.org/10.1016/j.rse.2017.09.036

Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., Nauss, T., 2018. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. Environmental Modelling & Software 101, 1–9. https://doi.org/10.1016/j.envsoft.2017.12.001

Navarro, L.M., Fernández, N., Guerra, C., Guralnick, R., Kissling, W.D., Londoño, M.C., Muller-Karger, F., Turak, E., Balvanera, P., Costello, M.J., Delavaud, A., El Serafy, G., Ferrier, S., Geijzendorffer, I., Geller, G.N., Jetz, W., Kim, E.-S., Kim, H., Martin, C.S., McGeoch, M.A., Mwampamba, T.H., Nel, J.L., Nicholson, E., Pettorelli, N., Schaepman, M.E., Skidmore, A., Sousa Pinto, I., Vergara, S., Vihervaara, P., Xu, H., Yahara, T., Gill, M., Pereira, H.M., 2017. Monitoring biodiversity change through effective global coordination. Current Opinion in Environmental Sustainability 29, 158–169. https://doi.org/10.1016/j.cosust.2018.02.005

Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Borman, T., Carvalho, G., Chirico, M., Caceres, M.D., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M.-H., Cunha, E.R., Smith, T., Stier, A., Braak, C.J.F.T., Weedon, J., 2025. vegan: Community Ecology Package.

Pinto-Ledezma, J.N., Cavender-Bares, J., 2020. Using remote sensing for modeling and monitoring species distributions, in: Cavender-Bares, J., Gamon, J.A., Townsend, P.A. (Eds.), Remote Sensing of Plant Biodiversity. Springer International Publishing, Cham, pp. 199–223. https://doi.org/10.1007/978-3-030-33157-3_9

Pinto-Ledezma, J.N., Larkin, D.J., Cavender-Bares, J., 2018. Patterns of beta diversity of vascular plants and their correspondence with biome boundaries across North America. Front. Ecol. Evol. 6, 194. https://doi.org/10.3389/fevo.2018.00194

Pinto-Ledezma, J.N., Schweiger, A.K., Guzmán Q., J.A., Cavender-Bares, J., 2025. Plant diversity across dimensions: Coupling biodiversity measures from the ground and the sky. Sci. Adv. 11, eadr0278. https://doi.org/10.1126/sciadv.adr0278

Qian, H., Swenson, N.G., Zhang, J., 2013. Phylogenetic beta diversity of angiosperms in N orth A merica. Global Ecology and Biogeography 22, 1152–1161. https://doi.org/10.1111/geb.12076

R Core Team, 2023. R: A Language and Environment for Statistical Computing.

Rivard, B., Feng, J., Gallie, A., Sanchez-Azofeifa, A., 2008. Continuous wavelets for the improved use of spectral libraries and hyperspectral data. Remote Sensing of Environment 112, 2850–2862. https://doi.org/10.1016/j.rse.2008.01.016

Rossi, C., Kneubühler, M., Schütz, M., Schaepman, M.E., Haller, R.M., Risch, A.C., 2022. Spatial resolution, spectral metrics and biomass are key aspects in estimating plant species richness from spectral diversity in species-rich grasslands. Remote Sens Ecol Conserv 8, 297–314. https://doi.org/10.1002/rse2.244

Ruefenacht, B., Finco, M.V., Nelson, M.D., Czaplewski, R., Helmer, E.H., Blackard, J.A., Holden, G.R., Lister, A.J., Salajanu, D., Weyermann, D., Winterberger, K., 2008. Conterminous U.S. and alaska forest type mapping using Forest Inventory and Analysis data. Photogrammetric Engineering & Remote Sensing 74, 1379–1388. https://doi.org/10.14358/PERS.74.11.1379

Schweiger, A.K., Cavender-Bares, J., Townsend, P.A., Hobbie, S.E., Madritch, M.D., Wang, R., Tilman, D., Gamon, J.A., 2018. Plant spectral diversity integrates functional and

phylogenetic components of biodiversity and predicts ecosystem function. Nat Ecol Evol 2, 976–982. https://doi.org/10.1038/s41559-018-0551-1
Schweiger, A.K., Laliberté, E., 2022. Plant beta-diversity across biomes captured by imaging spectroscopy. Nat Commun 13, 2767. https://doi.org/10.1038/s41467-022-30369-6
Silvertown, J., 2004. Plant coexistence and the niche. Trends in Ecology & Evolution 19, 605–611. https://doi.org/10.1016/j.tree.2004.09.003
Skidmore, A.K., Coops, N.C., Neinavaz, E., Ali, A., Schaepman, M.E., Paganini, M., Kissling, W.D., Vihervaara, P., Darvishzadeh, R., Feilhauer, H., Fernandez, M., Fernández, N., Gorelick, N., Geijzendorffer, I., Heiden, U., Heurich, M., Hobern, D., Holzwarth, S., Muller-Karger, F.E., Van De Kerchove, R., Lausch, A., Leitão, P.J., Lock, M.C., Mücher, C.A., O'Connor, B., Rocchini, D., Roeoesli, C., Turner, W., Vis, J.K., Wang, T., Wegmann, M., Wingate, V., 2021. Priority list of biodiversity metrics to observe from space. Nat Ecol Evol 5, 896–906. https://doi.org/10.1038/s41559-021-01451-x
Smith, S.A., Brown, J.W., 2018. Constructing a broadly inclusive seed plant phylogeny. American J of Botany 105, 302–314. https://doi.org/10.1002/ajb2.1019
Thompson, D.R., Green, R.O., Bradley, C., Brodrick, P.G., Mahowald, N., Dor, E.B., Bennett, M., Bernas, M., Carmon, N., Chadwick, K.D., Clark, R.N., Coleman, R.W., Cox, E., Diaz, E., Eastwood, M.L., Eckert, R., Ehlmann, B.L., Ginoux, P., Ageitos, M.G., Grant, K., Guanter, L., Pearlshtien, D.H., Helmlinger, M., Herzog, H., Hoefen, T., Huang, Y., Keebler, A., Kalashnikova, O., Keymeulen, D., Kokaly, R., Klose, M., Li, L., Lundeen, S.R., Meyer, J., Middleton, E., Miller, R.L., Mouroulis, P., Oaida, B., Obiso, V., Ochoa, F., Olson-Duvall, W., Okin, G.S., Painter, T.H., Pérez García-Pando, C., Pollock, R., Realmuto, V., Shaw, L., Sullivan, P., Swayze, G., Thingvold, E., Thorpe, A.K., Vannan, S., Villarreal, C., Ung, C., Wilson, D.W., Zandbergen, S., 2024. On-orbit calibration and performance of the EMIT imaging spectrometer. Remote Sensing of Environment 303, 113986. https://doi.org/10.1016/j.rse.2023.113986
Turner, W., 2014. Sensing biodiversity. Science 346, 301–302. https://doi.org/10.1126/science.1256014
Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A.K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L.P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M., Woodcock, C., 2015. Free and open-access satellite data are key to biodiversity conservation. Biological Conservation 182, 173–176. https://doi.org/10.1016/j.biocon.2014.11.048
USDA Forest Service, 2023. Forest Inventory and Analysis Database.
Wallis, C.I.B., Kothari, S., Jantzen, J.R., Crofts, A.L., St-Jean, S., Inamdar, D., Pablo Arroyo-Mora, J., Kalacska, M., Bruneau, A., Coops, N.C., Laliberté, E., Vellend, M., 2024. Exploring the spectral variation hypothesis for α- and β-diversity: a comparison of open vegetation and forests. Environ. Res. Lett. 19, 064005. https://doi.org/10.1088/1748-9326/ad44b1
Wang, R., Gamon, J.A., Schweiger, A.K., Cavender-Bares, J., Townsend, P.A., Zygielbaum, A.I., Kothari, S., 2018. Influence of species richness, evenness, and composition on optical diversity: A simulation study. Remote Sensing of Environment 211, 218–228. https://doi.org/10.1016/j.rse.2018.04.010
Wang, Z., Chlus, A., Geygan, R., Ye, Z., Zheng, T., Singh, A., Couture, J.J., Cavender-Bares, J., Kruger, E.L., Townsend, P.A., 2020. Foliar functional traits from imaging spectroscopy across biomes in eastern North America. New Phytologist 228, 494–511. https://doi.org/10.1111/nph.16711
Westfall, J.A., Schroeder, T.A., McCollum, J.M., Patterson, P.L., 2022. A spatial and temporal assessment of nonresponse in the national forest inventory of the U.S. Environ Monit Assess 194, 530. https://doi.org/10.1007/s10661-022-10219-0

Zurell, D., Bocedi, G., Velazco, S.J.E., Gonzalez, A., Purvis, A., Wintle, B., Merow, C., Lundquist, C., Guillera-Arroita, G., Settele, J., Serra-Diaz, J.M., Cabral, J.S., Travis, J.M.J., Schifferle, K., Buckley, L., Briscoe, N.J., Isaac, N.J.B., Peres-Neto, P.R., Keuth, R., Gascoigne, S.J.L., Ferrier, S., Urban, M.C., 2025. Predicting the way forward for the Global Biodiversity Framework. Proc. Natl. Acad. Sci. U.S.A. 122, e2501695122. https://doi.org/10.1073/pnas.2501695122

Figures and Tables

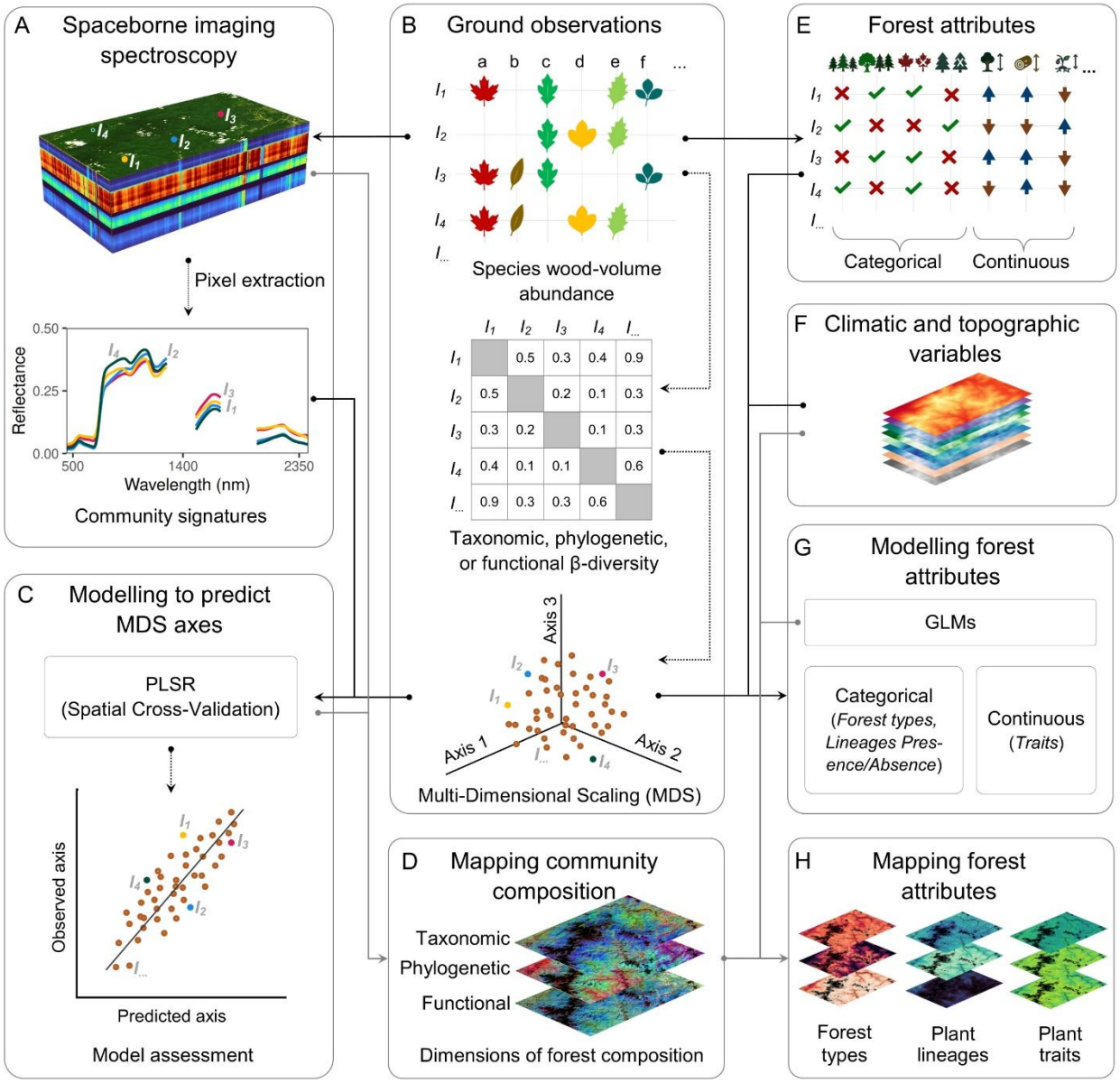


Fig. 1. Framework for mapping forest community composition and its attributes using spaceborne imaging spectroscopy. Forest inventory locations are first used to extract pixels from scenes (A). Species abundance data from inventories are then used to calculate pairwise β-diversity matrices (B), which are then ordained using multi-dimensional scaling (MDS). The resulting axes of β-diversity are combined with spectral information to model (C) and map forest community composition (D). These axes are also paired with forest attributes (E) and climatic and topographic variables (F) to model (G) and map a range of forest types, plant lineages, and community plant traits (H). Together, this framework enables the spatial mapping of both forest community composition and associated forest attributes at large spatial extent. Black solid arrows represent the link between data sources for modelling, grey solid arrows the use of data sources and models for mapping, and black dotted arrows intermediate steps.

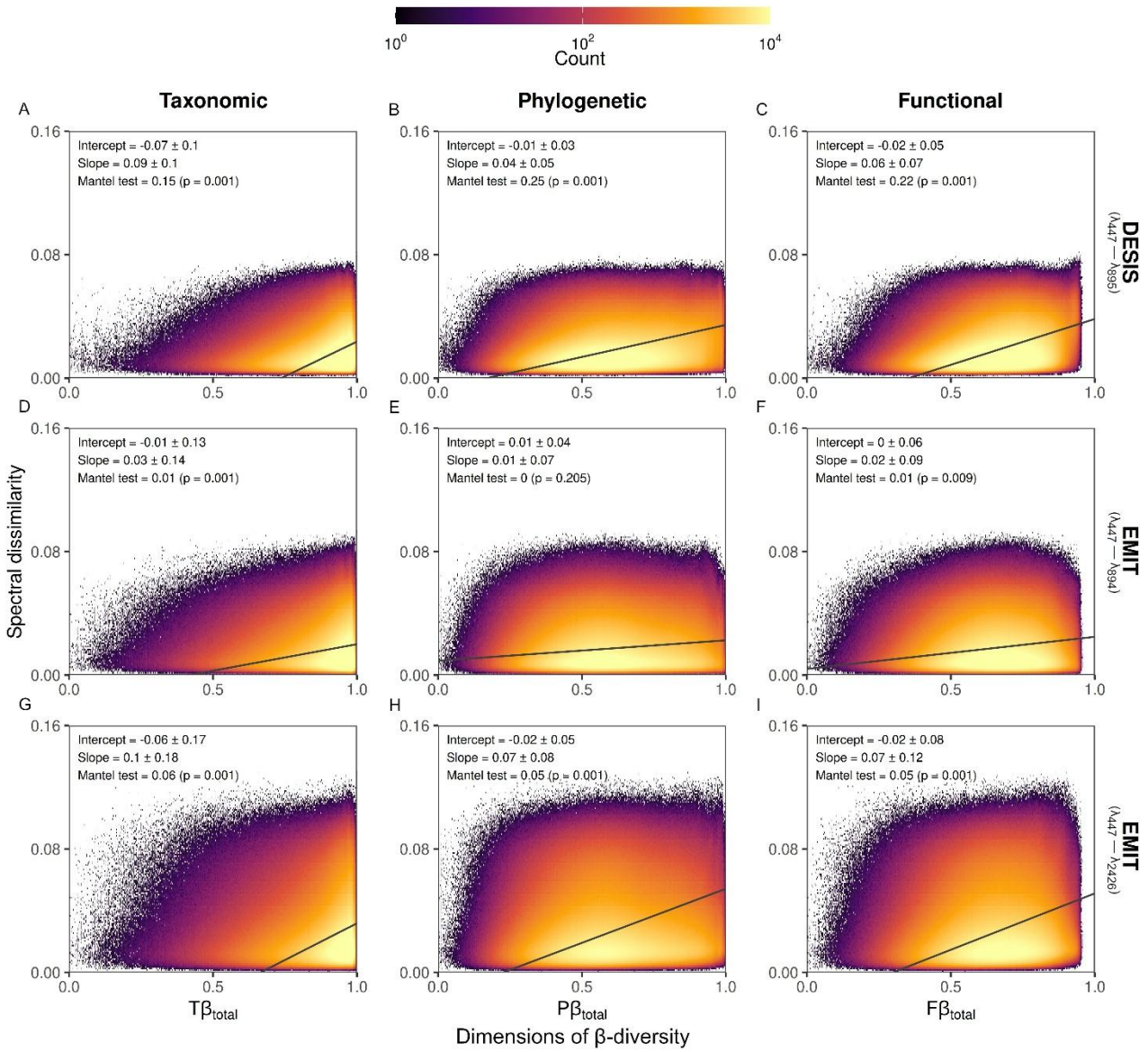


Fig. 2. Relationship between spectral dissimilarity and dimensions of β-diversity. Dimensions of beta total diversity (β_{total}) are represented by taxonomic ($T\beta_{total}$) (A, D, and G), phylogenetic ($P\beta_{total}$) (B, E, and H), and functional ($F\beta_{total}$) (C, F, and I) information. Spectral dissimilarly were estimated among communities from DESIS (A, B, and D) and EMIT observations, the latter using two spectral ranges: λ_{447}— λ_{895} (D, E, and F) and λ_{447} — λ_{2426} (G, H, and I). Black solid lines represent the average linear regression line when comparing dissimilarities from a target community with all communities. Extended figures are shown in Fig. S4–S5.

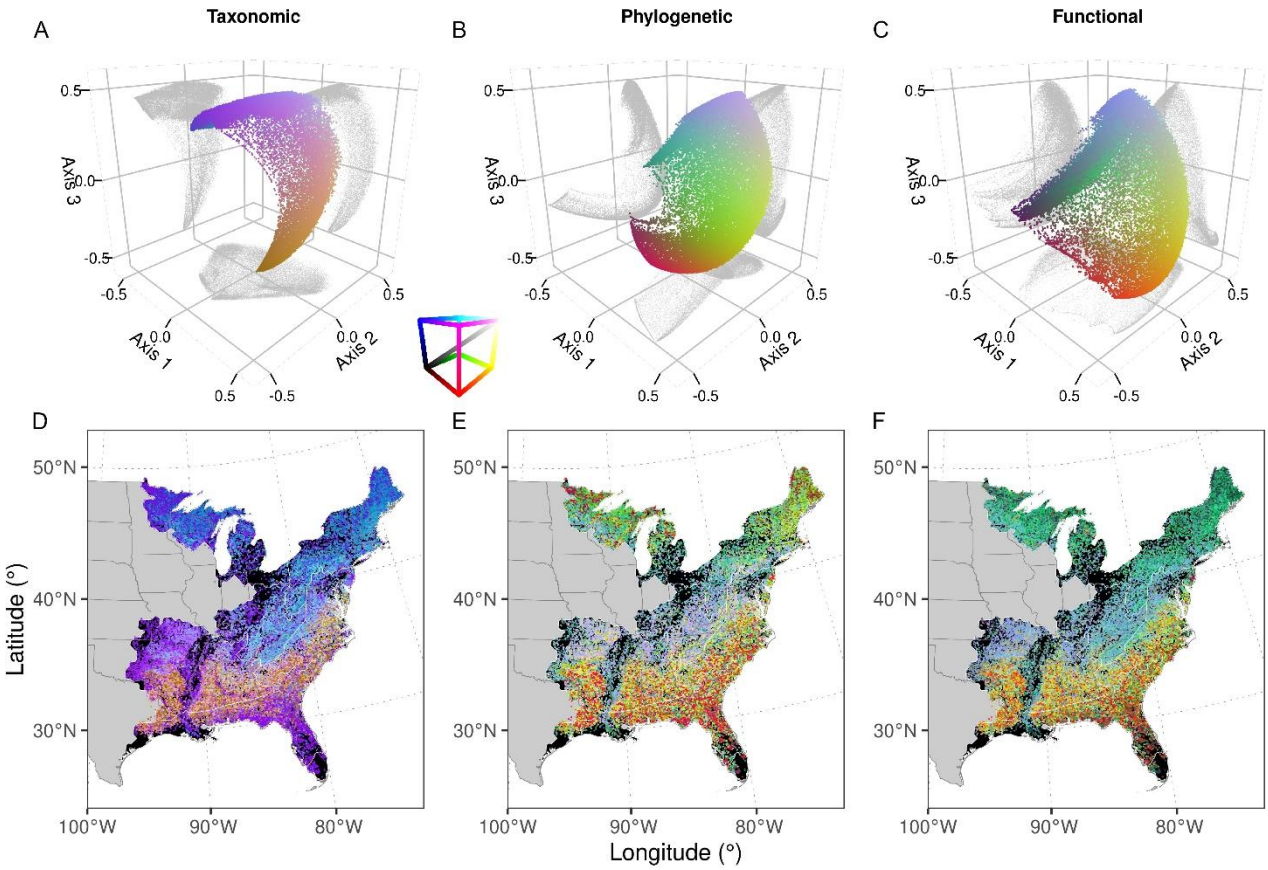


Fig. 3. Multi-Dimensional Scaling (MDS) used to ordinate dimensions of β-diversity. Panels shown ordinations of beta total diversity (β_{total}) and their spatial variation across the Eastern U.S. for taxonomic (A and D), phylogenetic (B and E), and functional (C and F) dimensions of diversity. Gray shading in panels A-C represents the projection of three-dimensional axes of β_{total} onto two-dimensional planes. Black areas in panels D, E, and F were included as a background to enhance the visual contrast of colors.

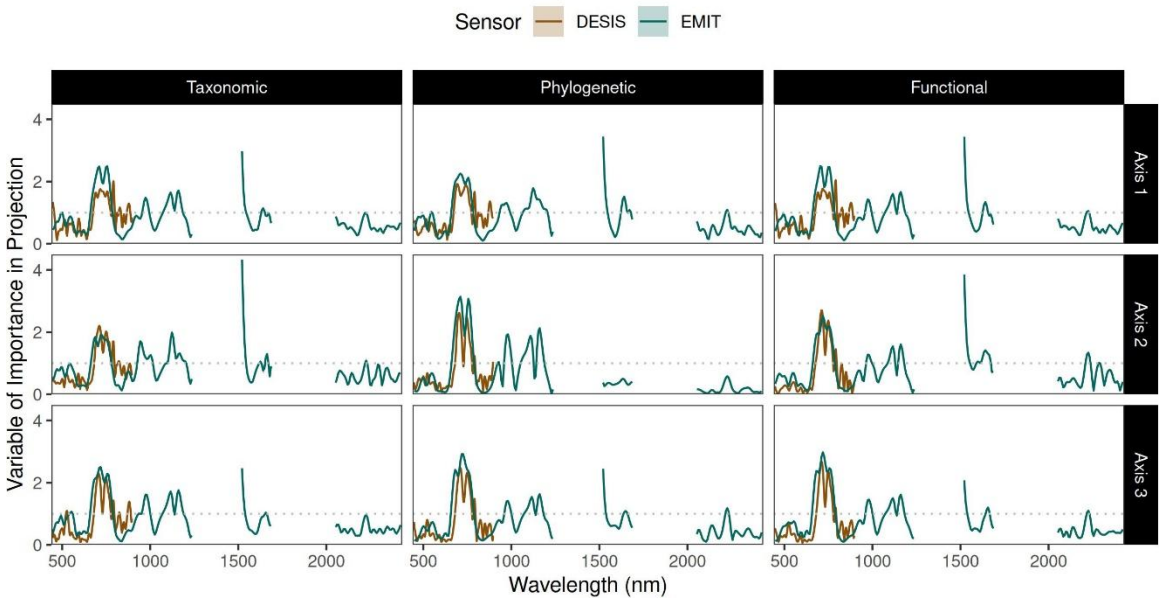


Fig. 4. Variable importance of projection across wavelengths of PLSR models to predict MDS axes of β-diversity. Lines represent the spaceborne observations of DESIS (brown) and EMIT (green) employed at different dimensions of diversity.

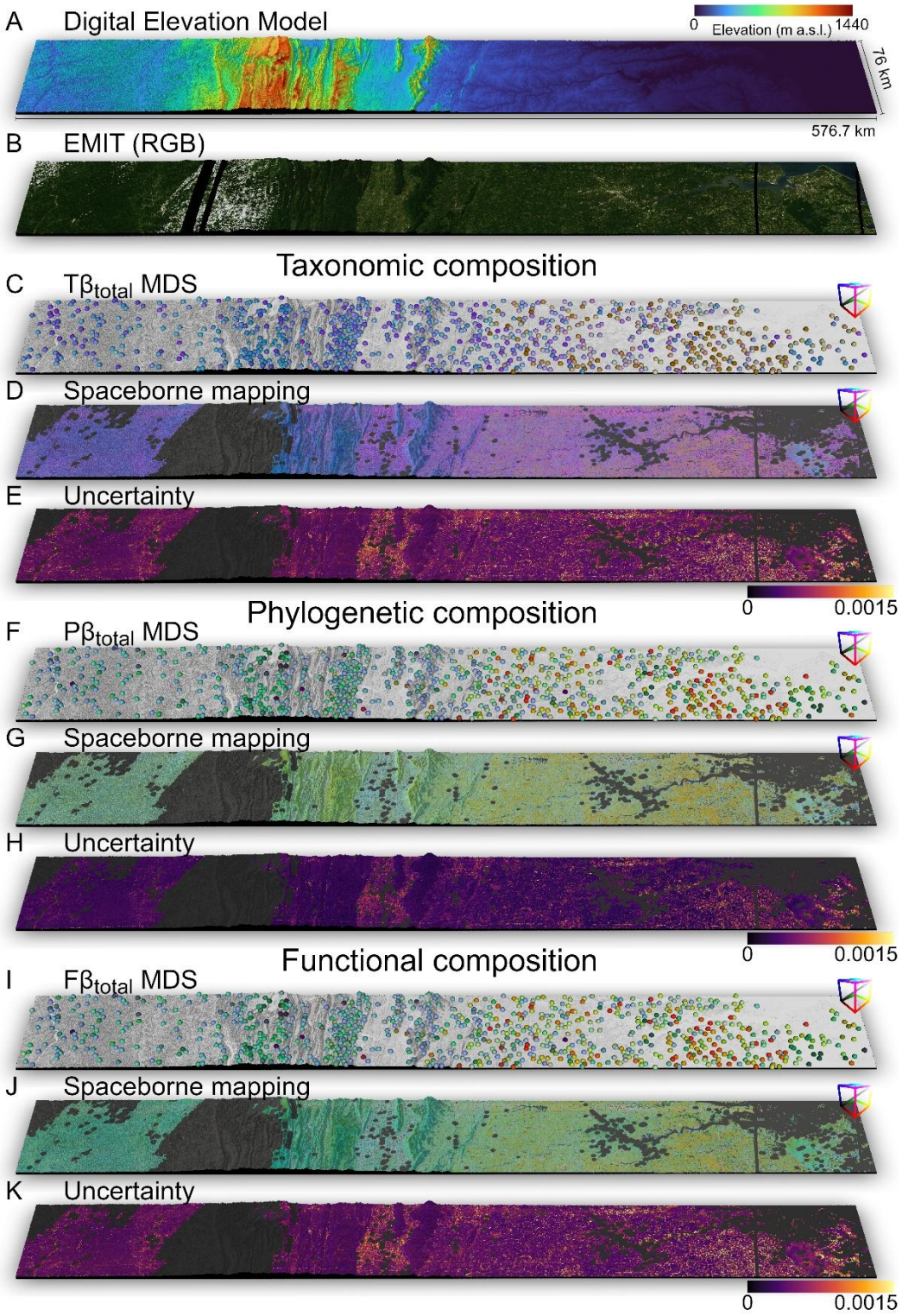


Fig. 5. Mapping multiple dimensions of forest community composition using EMIT across the Eastern United States. The first two panels show a true-color image (R: λ663, G: λ551, B: λ425) (A) and the provided elevation model with EMIT imagery (B). Panels C, F, and I represent dimensions of MDS β-diversity axes from plot inventories, while panels D, G, and J show their predicted spatial mapping from spaceborne observations and the associated uncertainty (E, H, and K). Each pixel value in the prediction and uncertainty maps represents an average of 100 repeated model runs.

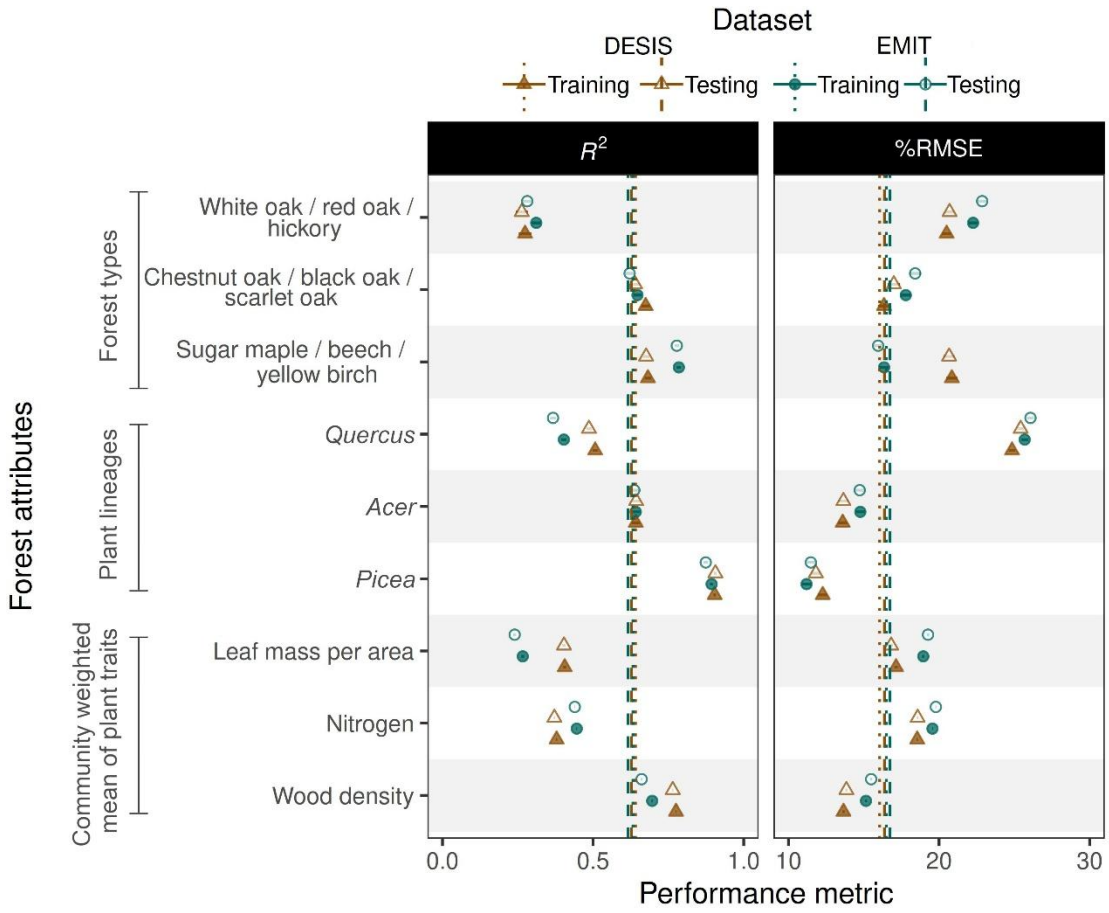


Fig. 6. Model performance for predicting attributes of forest composition from spaceborne observations from DESIS or EMIT. Each point represents an average of 100 repeated models. Vertical dashed lines represent the average performance for all forest attributes. Extended figures are shown in Fig. S30–S32.

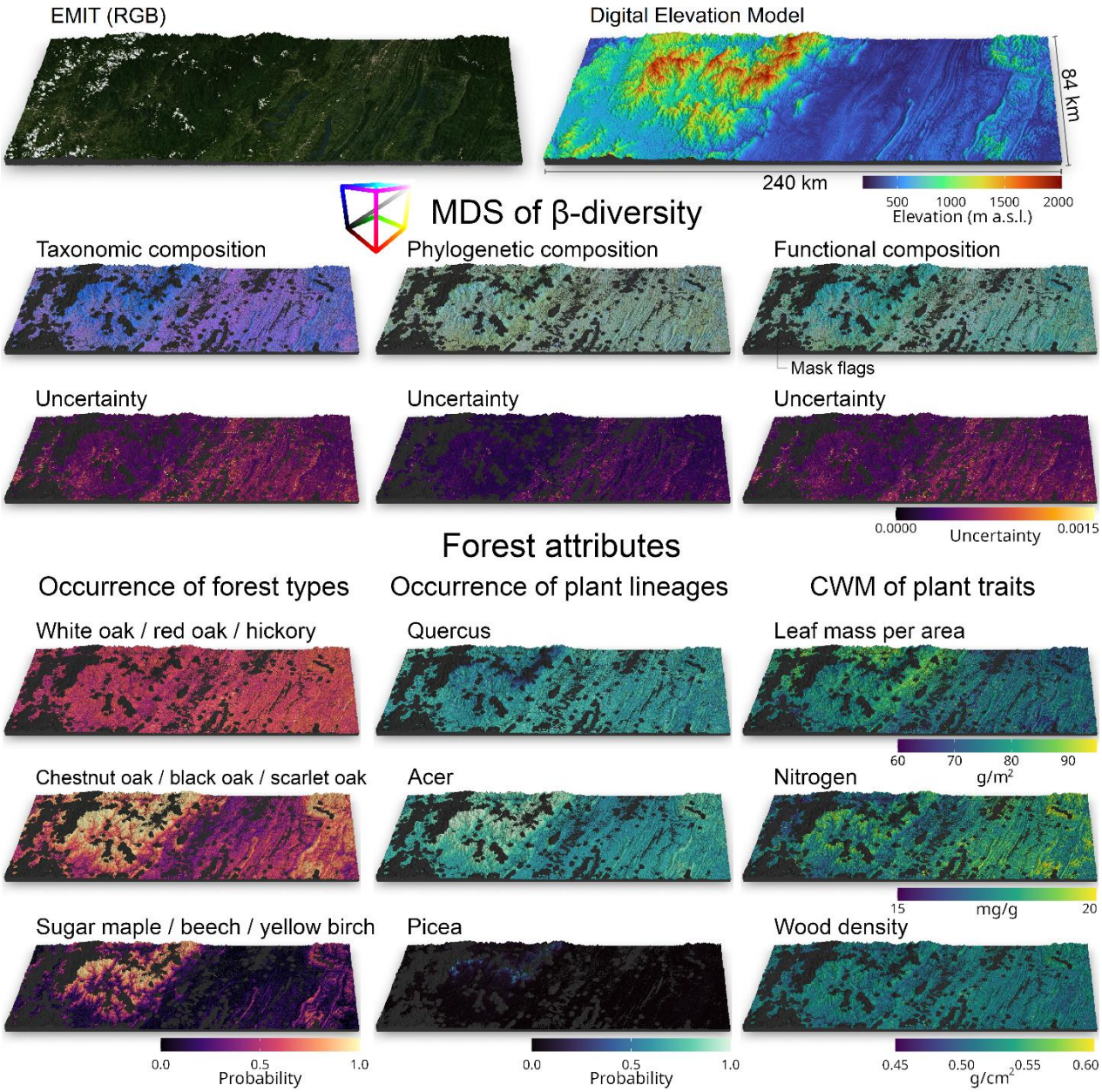


Fig 7. Mapping of multiple dimensions of forest diversity and their attributes using EMIT across the southeast of the United States. Top panels represent a true color image (R: λ_{663}, G: λ_{551}, B: λ_{425}) and the elevation model provided by NASA with EMIT imagery. Middle panels of community composition represent the mapping of MDS β-diversity axes based on different dimensions of diversity as well as their uncertainty. Bottom panels describe the mapping of forest attributes associated with the occurrence of forest types, plant lineages, and CWM of plant traits using the predicted MDS β-diversity axes, climatic, and topographic variables. Each pixel value represents an average of 100 repeated models.

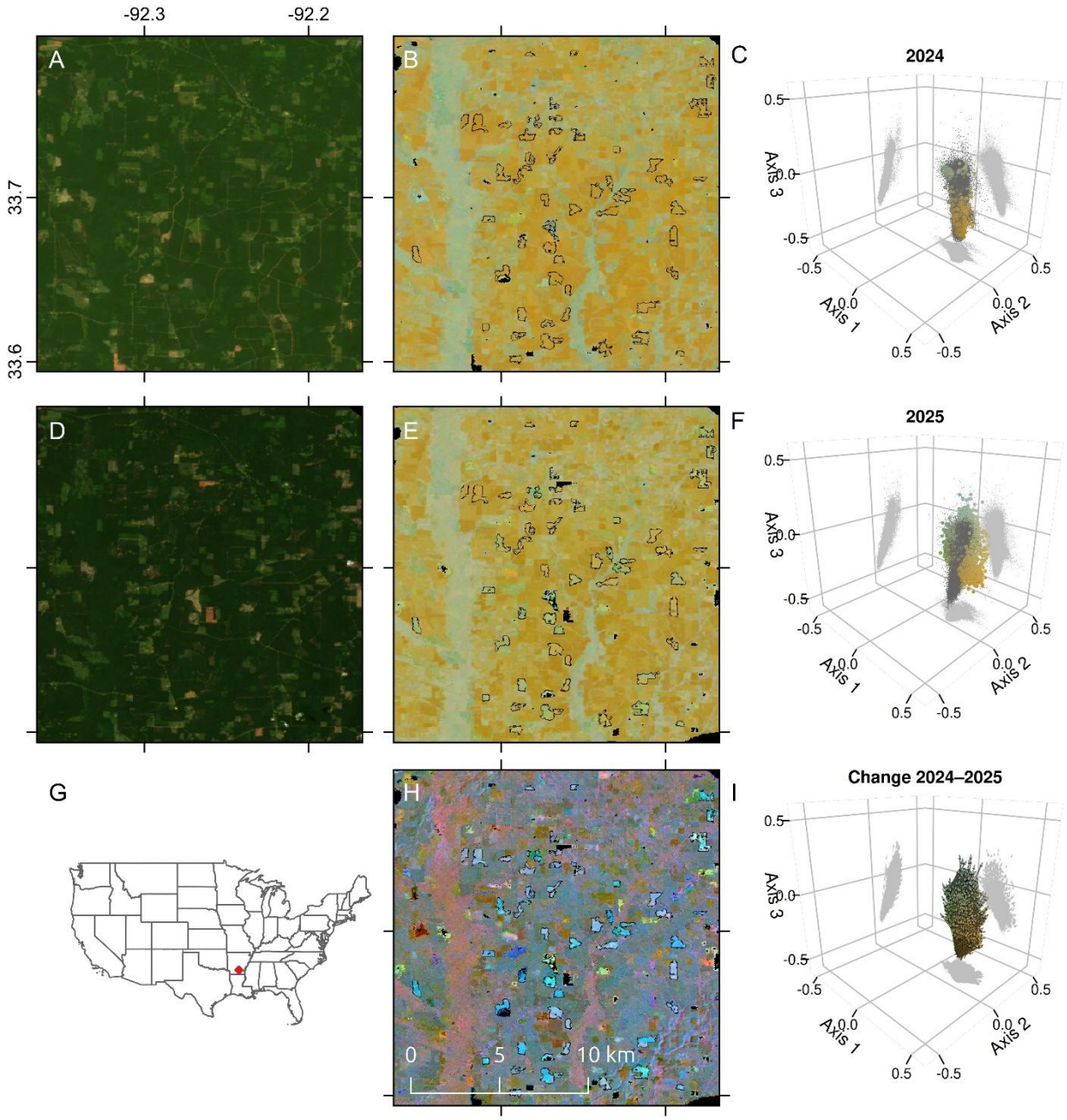


Fig. 8. Spatiotemporal variation in forest community composition derived from EMIT in Arkansas. True-color images (R: λ_{663}, G: λ_{551}, B: λ_{425}) are shown in panels A and D for 2024 and 2025, respectively. Predicted maps of forest community composition based on MDS Pβ-diversity axes are shown in panels B and E for 2024 and 2025. Panel H illustrates the spatiotemporal differences between MDS Pβ-diversity axes from 2024 to 2025, highlighting areas of change in forest community composition. Panels A, B, D, E, and H include outlined polygons indicating areas with major changes in community composition. The distribution of pixels in MDS ordination space is shown in panels C and F for 2024 and 2025; colored points correspond to pixels located within the polygons, while gray points represent the full regional distribution. Corresponding shifts in ordination space between 2024 and 2025 are shown in panel I, highlighting directional changes in community composition for pixels within the selected polygons. Panel G shows the location of the study area within the United States.

Supplementary Materials for

Mapping multiple dimensions of forest diversity using spaceborne spectroscopy

J. Antonio Guzmán Q.[1,2]*, Jonathan A. Knott [3], Jesús N. Pinto-Ledezma [2], Philip A. Townsend [4], Jeannine Cavender-Bares [1,2]*

[1] Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA.
[2] Department of Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, MN 55108, USA.
[3] USDA Forest Service, Northern Research Station, Saint Paul, MN 55108, USA.
[4] Department of Forest and Wildlife Ecology, University of Wisconsin—Madison, Madison, WI 53706, USA.

*Corresponding authors: aguzman@fas.harvard.edu and jcavender@fas.harvard.edu

This PDF file includes:

Methods S1. Selection of inventoried plots from the Forest Inventory and Analysis (FIA) program across the Eastern U.S.

The FIA program currently conducts inventories using an annualized inventory system, collecting data on a subset of plots each year, generally 1/7 to 1/5 of plots in the Eastern U.S., leading to a 5- or 7-year inventory cycle. We selected FIA plots from Eastern U.S. states by FIA evaluation ID (EVALID) with inventory cycles ending in 2018-2022, the most recent inventories available as of the data query on 27 April 2023 (Table S1). We limited our query to plots with at least one forested condition. From these plots, we selected all live trees > 12.7 cm diameter at breast height (DBH) that are sampled on four circular 7.31 m radius subplots. Tree diameter and height were used to estimate wood volume assuming a cylinder shape, and volume was converted to abundance per unit area by multiplying by FIA trees-per-acre adjustment factor (TPA_UNADJ), which is based on the ratio of the subplot sampling area to one acre of land. We used the National Ecological Observatory Network (NEON) ecoregion map (Hargrove and Hoffman, 1999) to further filter plots from forested ecoregions.

Methods S2. Integration of plant functional traits.

We collected data for eight plant functional traits: tree maximum height (H), tree slenderness (S), wood density (WD), shade tolerance index (STI), drought tolerance index (DTI), waterlogging tolerance index (WTI), leaf mass per area (LMA), and nitrogen content (N). H (m) was estimated as the 95-quantile value of tree height per species among all the FIA plots that were used. S was estimated as the 50-quantile value of the ratio between tree height and diameter at the breast height for all the trees per species among all the FIA plots that were used. Wood density (g m^{-3}) was obtained directly from the FIA database per each species. STI, DTI, and WTI were obtained from (Niinemets and Valladares, 2006). LMA (g m^{-2}) and N (mg g^{-1}) were obtained from the Botanical Information and Ecology Network (BIEN) database through R (Maitner et al., 2018), accessed on April 6, 2024. Data were missing for some species for traits gathered from (Niinemets and Valladares, 2006) and BIEN, and therefore, we imputed them using a phylogenetic approach following (Debastiani et al., 2021). We used a pruned phylogenetic tree from (Smith and Brown, 2018) to generate a phylogenetic distance matrix of ten orthogonal eigenvectors using the 'PVRdecomp' function of the PVR package of R (Santos, 2018). All plant traits with missing values and the ten orthogonal eigenvectors were integrated into the 'missForest' function (Stekhoven and Bühlmann, 2012) using 500 iterations and 100 trees to impute missing values. The resulting imputed values were used to fill in missing traits (e.g., Fig. S31).

We applied a principal components analysis (PCA) to the complete set of plant traits per species to summarize the variation of traits among species into three principal components. The first, second, and third principal components explained 24.28, 17.88, and 16.32% of the variation in plant traits (Fig. S32), respectively. We then extracted PCA scores of the first three principal components per species to construct a functional trait dendrogram using euclidean distances among species. Finally, the functional trait dendrogram was transformed into a phylogenetic tree class in R using the "as.phylogeny" command in APE (Paradis et al., 2004) to generate functional dissimilarity values for the calculation of functional beta diversity.

Methods S3. Estimation of spectral dissimilarity and its regression with dimensions of beta diversity.

We compared pairwise spectral dissimilarity values among communities on selected clear-sky spaceborne observations of FIA plots. These pairwise comparisons were calculated using smoothed reflectance spectra, not on wavelet transformations of spectra. To smooth the reflectance spectra, we applied a Savitzky-Golay smoothing filter on the extracted pixels from the FIA plots. This filter was applied using a first order and two neighboring bands of length from the target bands (i.e., $n = 5$). The application of the Savitzky-Golay filter was done using the 'sgolayfilt' function of the signal package in R (signal developers, 2023). We then removed bands on the resulting smoothed spectra to match spectral regions used in our modelling efforts (ie., DESIS: 449 – 965 nm; EMIT: 447 – 1265, 1488 – 1710, and 2015 – 2424 nm).

Methods S4. Climatic and topographic datasets.

We used well-established climatic and topographic datasets at 1 km of spatial resolution to rotate MDS β-diversity axes and predict forest attributes. Specifically, we employed historic climatic datasets for mean annual temperature (MAT), temperature seasonality (TS), temperature annual range (TAR), mean annual precipitation (MAP), and precipitation seasonality (PS) derived from Worldclim 2 (Fick and Hijmans, 2017). Likewise, we employed topographic datasets associated with the elevation (ELE), slope (SL), and compound topographic index (CTI) (e.g., wetness index) derived from HYDRO1k (EROS, 2017). The extraction of pixels from these datasets were conducted using publicly available coordinates of FIA plots given their coarse resolution. To integrate these climatic and topographic datasets with scenes of forest community composition to map forest attributes, we extracted values using the centroid location of pixels from spaceborne observations.

Table S1. Inventory end years, states, and FIA evaluation IDs (EVALIDs) included in this study.

Inventory end year	States	EVALIDs
2018	Kentucky	211801
2019	Florida, Louisiana, Massachusetts, New York, Ohio, Oklahoma, Tennessee	121901, 221901, 251901, 361901, 391901, 401901, 471901
2020	Kansas, Maryland, Michigan, Minnesota, New Jersey, Vermont, West Virginia	202001, 242001, 262001, 272001, 342001, 502001, 542001
2021	Arkansas, Connecticut, Delaware, Georgia, Illinois, Indiana, Maine, Mississippi, Missouri, New Hampshire, North Carolina, Pennsylvania, Rhode Island, South Carolina, Texas, Virginia, Wisconsin	52101, 92101, 102101, 132101, 172101, 182101, 232101, 282101, 292101, 332101, 372101, 422101, 442101, 452101, 482121, 512101, 552101
2022	Alabama	12201

Table S2. Performance summary of repeated Partial Least Square Regression models using spaceborne observations of DESIS to predict Multi-Dimensional Scaling (MDS) axes based on matrices of dimensions of beta diversity. Performance metrics are described by the coefficient of determination (R^2), bias, the Root Mean Square Error (RMSE), and the percentage of RMSE (%RMSE) based on the data range. Values represent the mean and standard deviation of 100 models.

Dimension	Dataset	Axes	Performance metric			
			R^2	BIAS (x10^{-5})	RMSE	%RMSE
Taxonomic	Training	Axis 1	0.38 ± 0.00	-6.9 ± 51.71	0.17 ± 0.00	21.84 ± 0.01
		Axis 2	0.14 ± 0.00	-3.3 ± 33.15	0.15 ± 0.00	25.29 ± 0.01
		Axis 3	0.08 ± 0.00	14.6 ± 20.27	0.12 ± 0.00	18.25 ± 0.00
	Testing	Axis 1	0.36 ± 0.00	402.7 ± 53.60	0.16 ± 0.00	21.87 ± 0.02
		Axis 2	0.11 ± 0.00	-320.6 ± 36.32	0.16 ± 0.00	25.92 ± 0.02
		Axis 3	0.07 ± 0.00	377.6 ± 23.62	0.12 ± 0.00	18.90 ± 0.01
Phylogenetic	Training	Axis 1	0.17 ± 0.00	2.2 ± 32.86	0.14 ± 0.00	20.70 ± 0.01
		Axis 2	0.20 ± 0.00	15 ± 43.89	0.19 ± 0.00	20.33 ± 0.01
		Axis 3	0.33 ± 0.00	-7.3 ± 72.08	0.26 ± 0.00	27.15 ± 0.01
	Testing	Axis 1	0.15 ± 0.00	-156.7 ± 36.13	0.14 ± 0.00	21.29 ± 0.02
		Axis 2	0.17 ± 0.00	631.3 ± 45.19	0.18 ± 0.00	20.33 ± 0.01
		Axis 3	0.31 ± 0.00	1265.3 ± 72.83	0.26 ± 0.00	27.36 ± 0.02
Functional	Training	Axis 1	0.30 ± 0.00	0.4 ± 48.43	0.19 ± 0.00	23.28 ± 0.01
		Axis 2	0.16 ± 0.00	19.3 ± 30.56	0.16 ± 0.00	20.03 ± 0.00
		Axis 3	0.30 ± 0.00	13.8 ± 60.67	0.20 ± 0.00	22.08 ± 0.01
	Testing	Axis 1	0.27 ± 0.00	112.3 ± 51.48	0.19 ± 0.00	23.31 ± 0.02
		Axis 2	0.13 ± 0.00	378 ± 31.04	0.15 ± 0.00	20.25 ± 0.01
		Axis 3	0.29 ± 0.00	731.1 ± 63.07	0.20 ± 0.00	22.51 ± 0.01

Table S3. Performance summary of repeated Partial Least Square Regression models using spaceborne observations of EMIT to predict Multi-Dimensional Scaling (MDS) axes based on matrices of dimensions of beta diversity. Performance metrics are described by the coefficient of determination (R^2), bias, the Root Mean Square Error (RMSE), and the percentage of RMSE (%RMSE) based on the data range. Values represent the mean and standard deviation of 100 models.

Dimension	Evaluation	Axes	Performance metric			
			R^2	BIAS ($x10^{-5}$)	RMSE	%RMSE
Taxonomic	Training	Axis 1	0.64 ± 0.00	-3.90 ± 58.41	0.13 ± 0.00	17.07 ± 0.01
		Axis 2	0.14 ± 0.00	-1.70 ± 28.04	0.14 ± 0.00	23.84 ± 0.01
		Axis 3	0.33 ± 0.00	32.50 ± 54.22	0.17 ± 0.00	18.88 ± 0.01
	Testing	Axis 1	0.61 ± 0.00	212.30 ± 59.32	0.13 ± 0.00	17.40 ± 0.02
		Axis 2	0.13 ± 0.00	-290.50 ± 31.28	0.15 ± 0.00	24.30 ± 0.01
		Axis 3	0.31 ± 0.00	96.80 ± 55.58	0.18 ± 0.00	19.60 ± 0.02
Phylogenetic	Training	Axis 1	0.27 ± 0.00	2.90 ± 34.71	0.12 ± 0.00	19.56 ± 0.01
		Axis 2	0.10 ± 0.00	20.40 ± 32.47	0.19 ± 0.00	21.24 ± 0.00
		Axis 3	0.27 ± 0.00	7.90 ± 74.02	0.29 ± 0.00	29.46 ± 0.01
	Testing	Axis 1	0.24 ± 0.00	-578.90 ± 36.46	0.13 ± 0.00	20.55 ± 0.02
		Axis 2	0.08 ± 0.00	349.00 ± 32.36	0.19 ± 0.00	20.67 ± 0.01
		Axis 3	0.24 ± 0.00	483.00 ± 77.06	0.29 ± 0.00	29.63 ± 0.02
Functional	Training	Axis 1	0.50 ± 0.00	-3.60 ± 59.64	0.17 ± 0.00	19.49 ± 0.01
		Axis 2	0.07 ± 0.00	13.40 ± 24.05	0.15 ± 0.00	20.55 ± 0.01
		Axis 3	0.31 ± 0.00	10.40 ± 65.01	0.23 ± 0.00	24.07 ± 0.01
	Testing	Axis 1	0.46 ± 0.00	261.90 ± 60.66	0.17 ± 0.00	20.15 ± 0.02
		Axis 2	0.05 ± 0.00	229.30 ± 25.63	0.15 ± 0.00	20.39 ± 0.02
		Axis 3	0.30 ± 0.00	32.10 ± 64.04	0.23 ± 0.00	24.47 ± 0.02

Table S4. Detailed list of spectral regions, its sensitivity, and potential ecological meaning with influence for predicting MDS β-diversity axes of multiple dimensions of diversity using spaceborne observations of DESIS and EMIT.

Wavelengths (nm)	Biophysical/biochemical sensitivity	Potential meaning associated with plant communities
553	Reflectance peak due to low absorption by chlorophyll	Proxy for variation in chlorophyll content and canopy health among communities
710	Chlorophyll absorption shoulder	Photosynthetic capacity and leaf nitrogen content among communities
720	Start of the near red-edge inflection	Sensitive to chlorophyll and canopy stress; early sign of senescence
756	Near red-edge inflection	Strong in vegetation indices; indicates species differences
790	Canopy structure	Sensitive to variations in leaf area index, canopy density, and leaf scattering properties among communities
973	Water absorption	Community-level water content
1116	Leaf structural traits and dry matter	Associated with dry matter, LMA, and some lignin/cellulose content
1161	Continuation of dry matter region	Tracks cell structure, leaf dry matter; functionally differentiates among communities
1521	Strong water absorption	Community-level water content
2226	Dry matter, protein, cellulose	Strong absorption by lignin, cellulose, protein; reflects C:N ratio, functional type, successional status

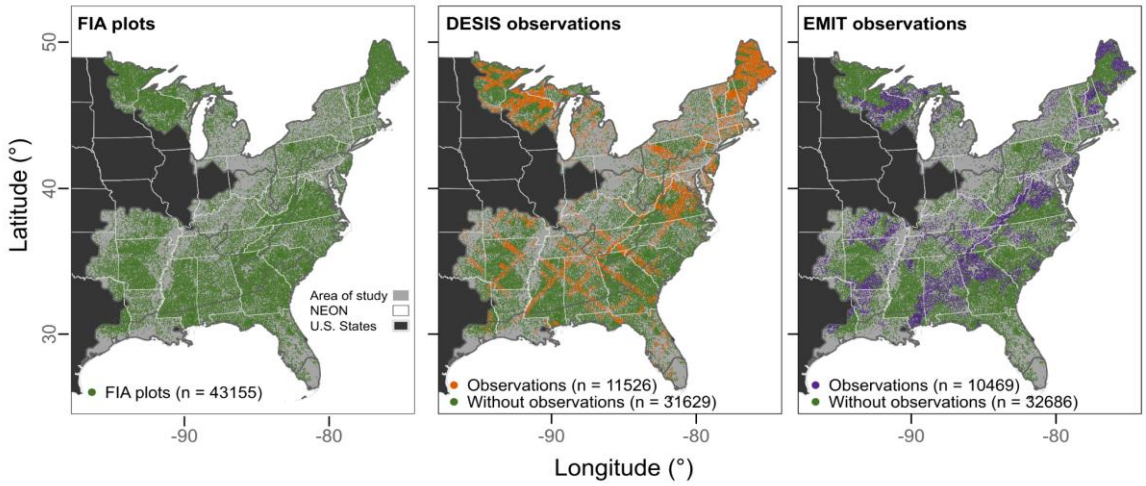


Fig. S1. Spatial distribution of perturbed plot locations from the Forest Inventory Analysis (FIA) program and the availability of clear sky observations from DESIS and EMIT spaceborne observations.

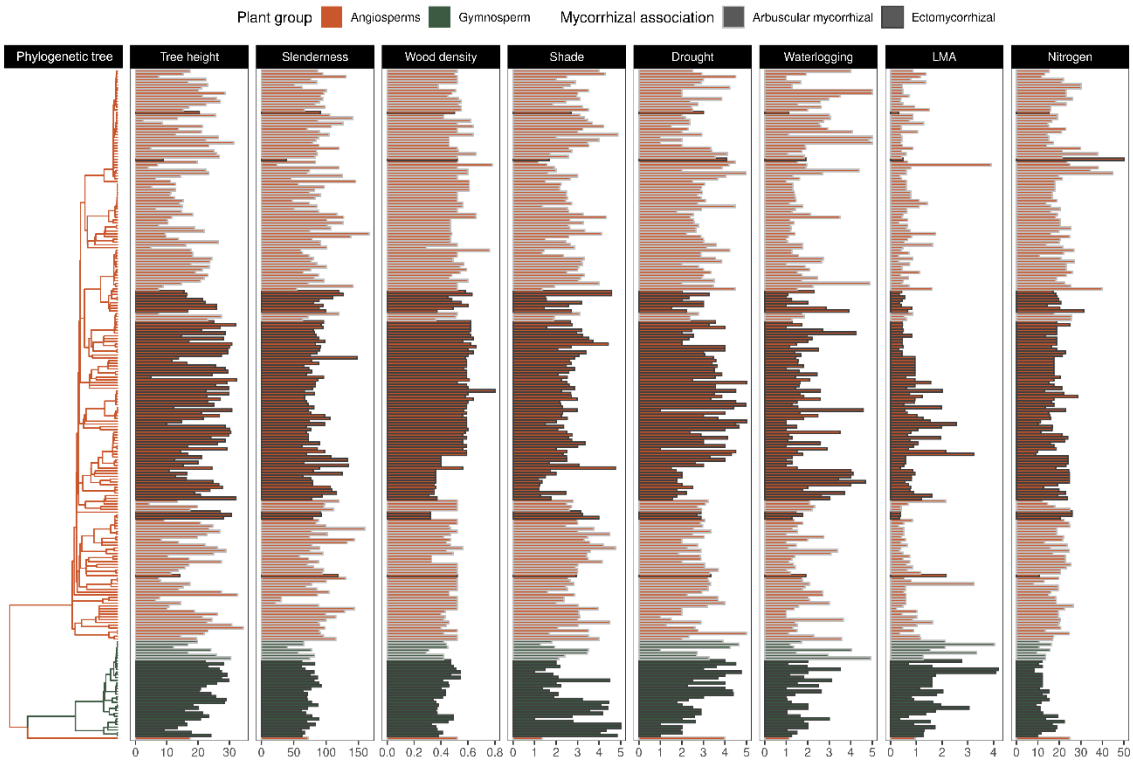


Fig. S2. Plant phylogenetic tree and its functional traits of inventoried species from the Forest Inventory and Analysis program across the Eastern U.S.

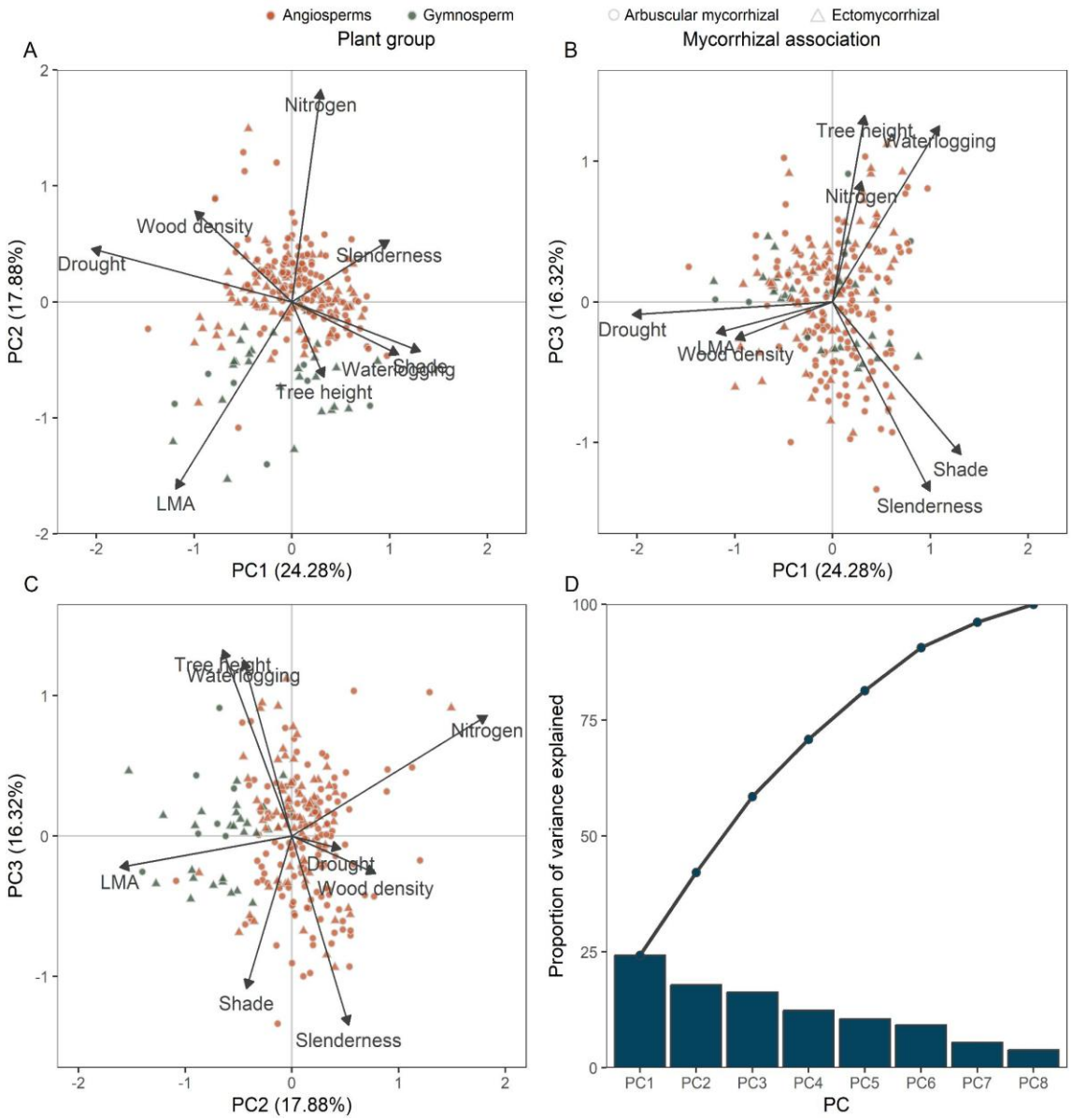


Fig. S3. Principal component analysis to summarize the variability of eight plant traits among species. Panels A, B, and C describe the biplot projections among the first three principal components, while panel D describes the proportion of variance explained by each component.

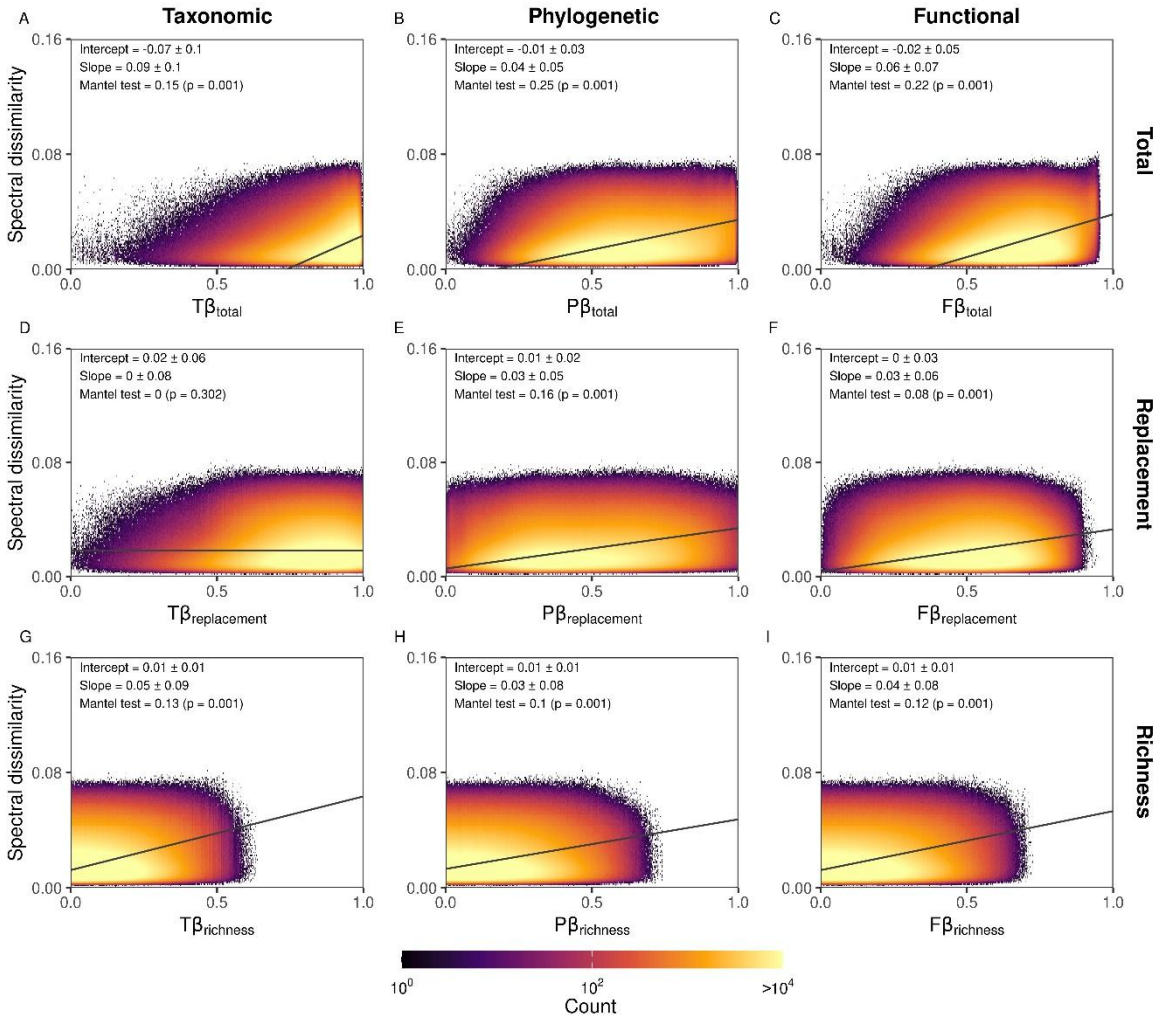


Fig. S4. Relationships between β- diversity (β) and spectral dissimilarity from DESIS observations among communities. Horizontal panels describe the dimensions of β diversity, including taxonomic (Tβ), phylogenetic (Pβ), and functional (Fβ) diversity. Vertical panels describe the partition of beta total diversity (β_{total}) into replacement ($\beta_{replacement}$) and richness ($\beta_{richness}$). The black solid lines represent the average linear regression line when regressing dissimilarities from a target community with all communities.

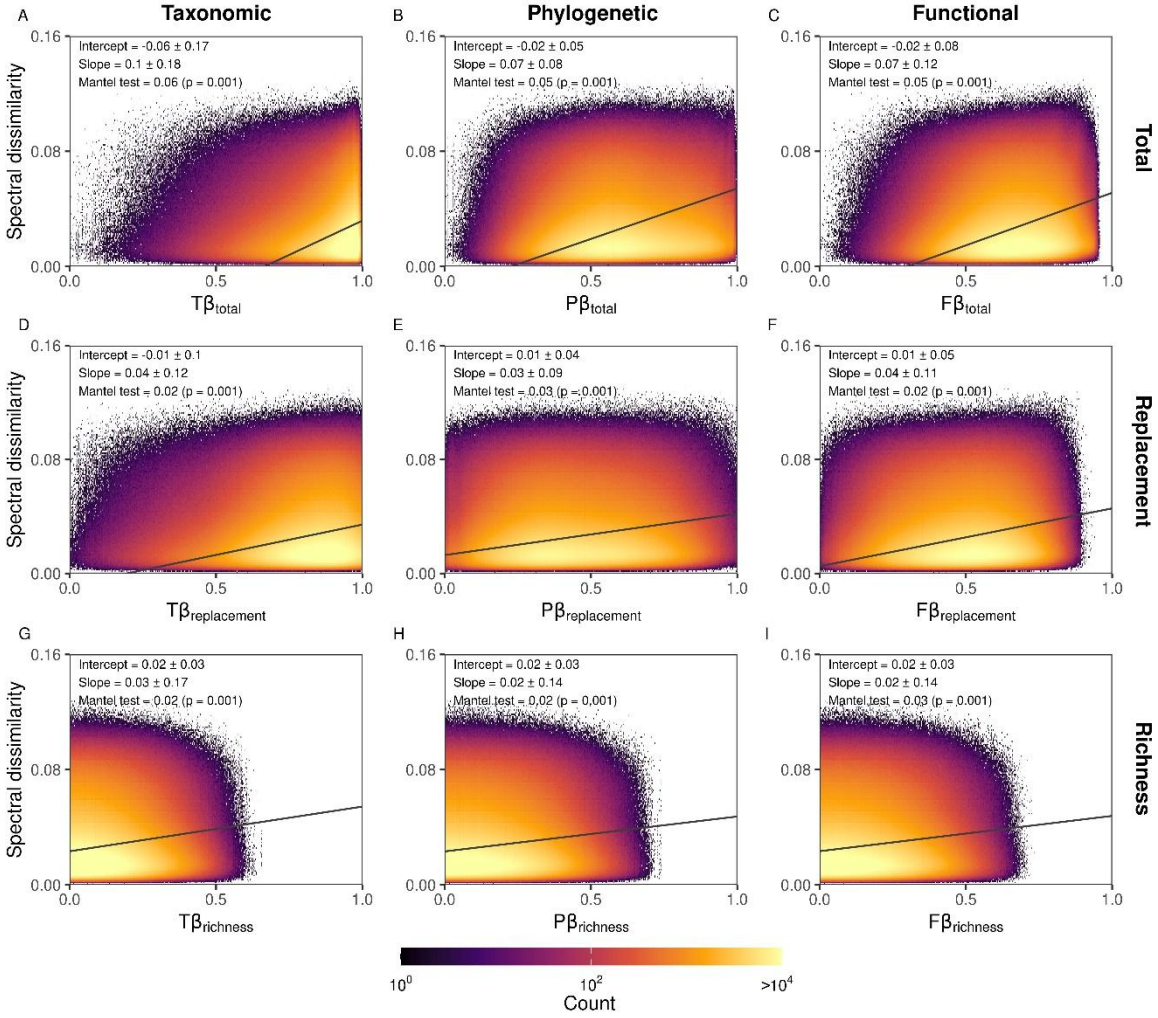


Fig. S5. Relationships between β-diversity (β) and spectral dissimilarity from EMIT observations among communities. Horizontal panels describe the dimensions of β diversity, including taxonomic (Tβ), phylogenetic (Pβ), and functional (Fβ) diversity. Vertical panels describe the partition of beta total diversity (β_{total}) into replacement ($\beta_{replacement}$) and richness ($\beta_{richness}$). The black solid lines represent the average linear regression line when regressing dissimilarities from a target community with all communities.

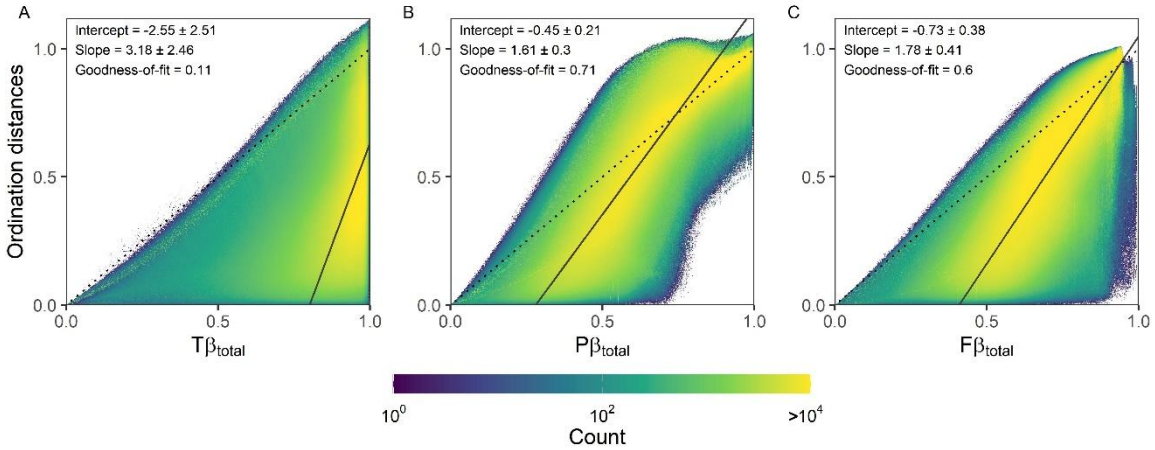


Fig. S6. Relationships between ordination distances derived from the Multi-Dimensional Scaling analyses and the dimensions of beta total diversity (β_{total}) based on taxonomic (A), phylogenetic (B), and functional (C) information. The black solid lines represent the average linear regression line when regressing distances between a target community and all other communities. Dotted lines represent the 1:1 line. The goodness-of-fit is calculated from the coefficient of determination of correlations between ordination axes values and β_{total} distances.

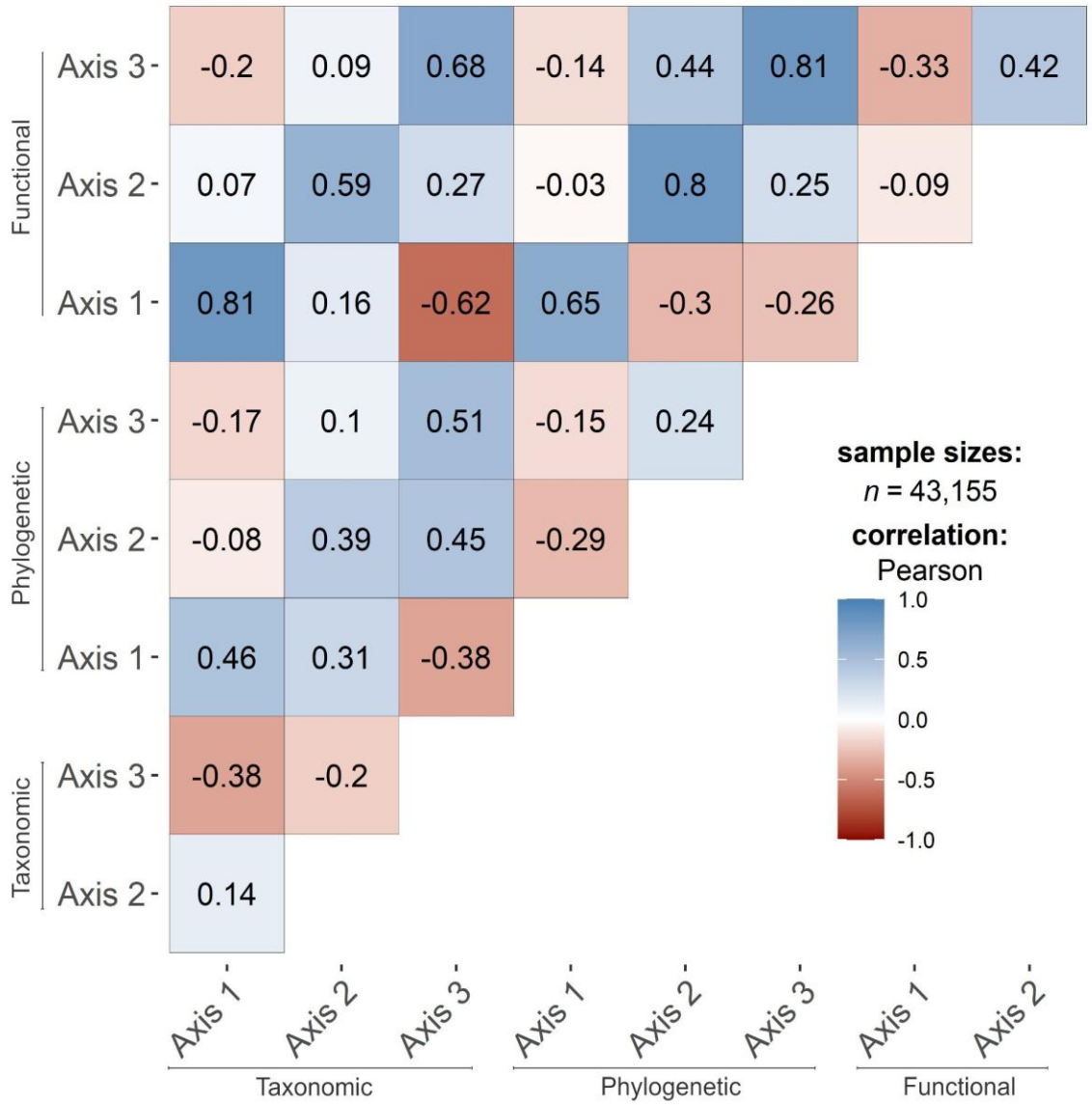


Fig. S7. Correlation coefficients between MDS axes of ordination for three dimensions of beta diversity, including taxonomic, phylogenetic, and functional β-diversity.

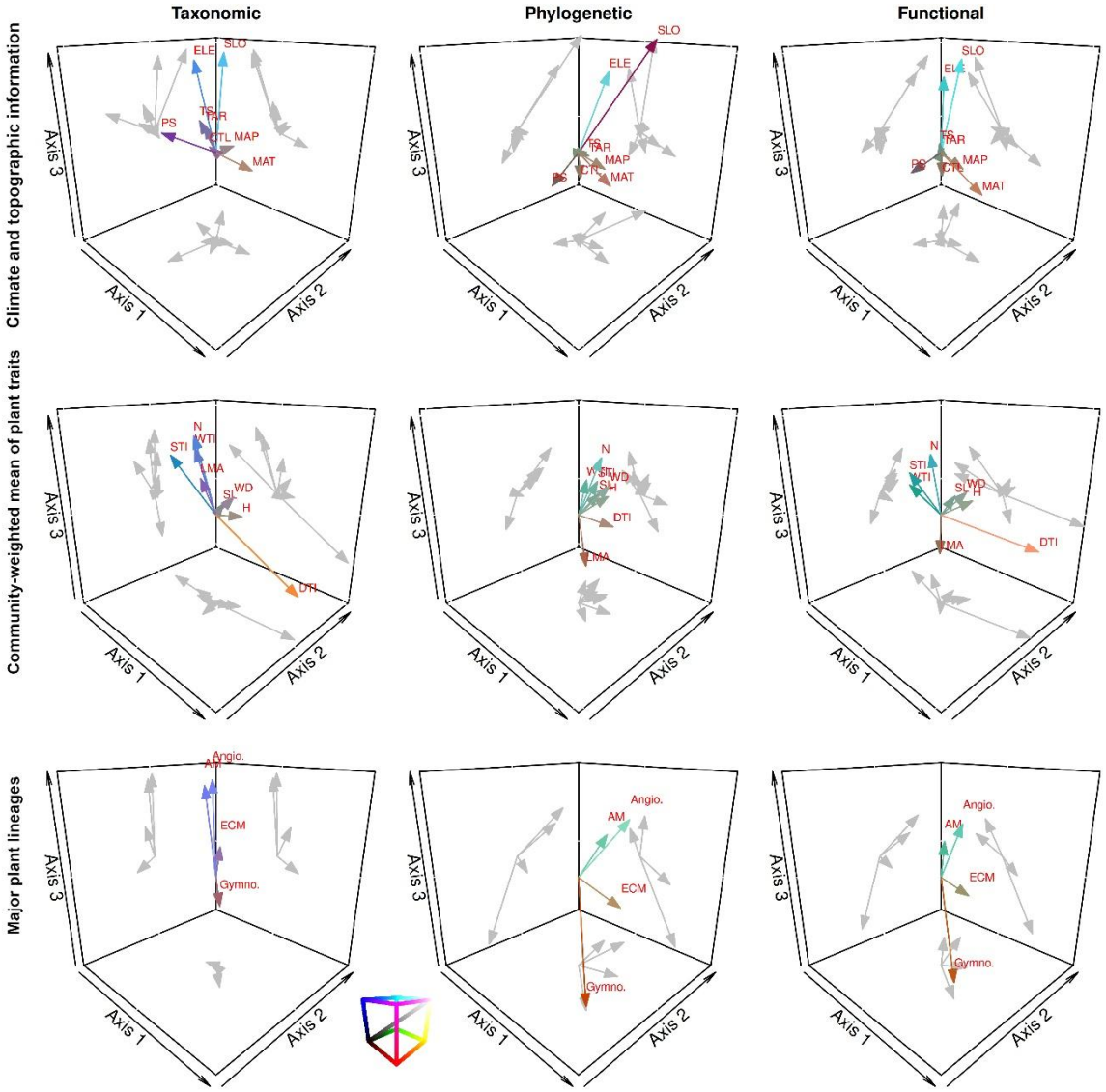


Fig. S8. Three-dimensional projection of weighted average scores. Vertical panels are arranged according to climatic and topography characteristics, community-weighted mean of plant traits, and relative abundance of major plant lineages of communities within Multi-Dimensional Scaling ordinations derived from taxonomic, phylogenetic, and functional β-diversity. Acronyms represent, mean annual temperature (MAT), temperature seasonality (TS), temperature annual range (TAR), annual precipitation (MAP), precipitation seasonality (PS), elevation (ELE), slope (SLO), and compound topographic index (CTI), tree maximum height (H), tree slenderness (SL), wood density (WD), shade tolerance index (STI), drought tolerance index (DTI), waterlogging tolerance index (WTI), leaf mass per area (LMA), nitrogen content (N), gymnosperms (Gymno.), angiosperms (Angio.), ectomycorrhizal symbiosis (ECM), arbuscular mycorrhizal symbiosis (AM).

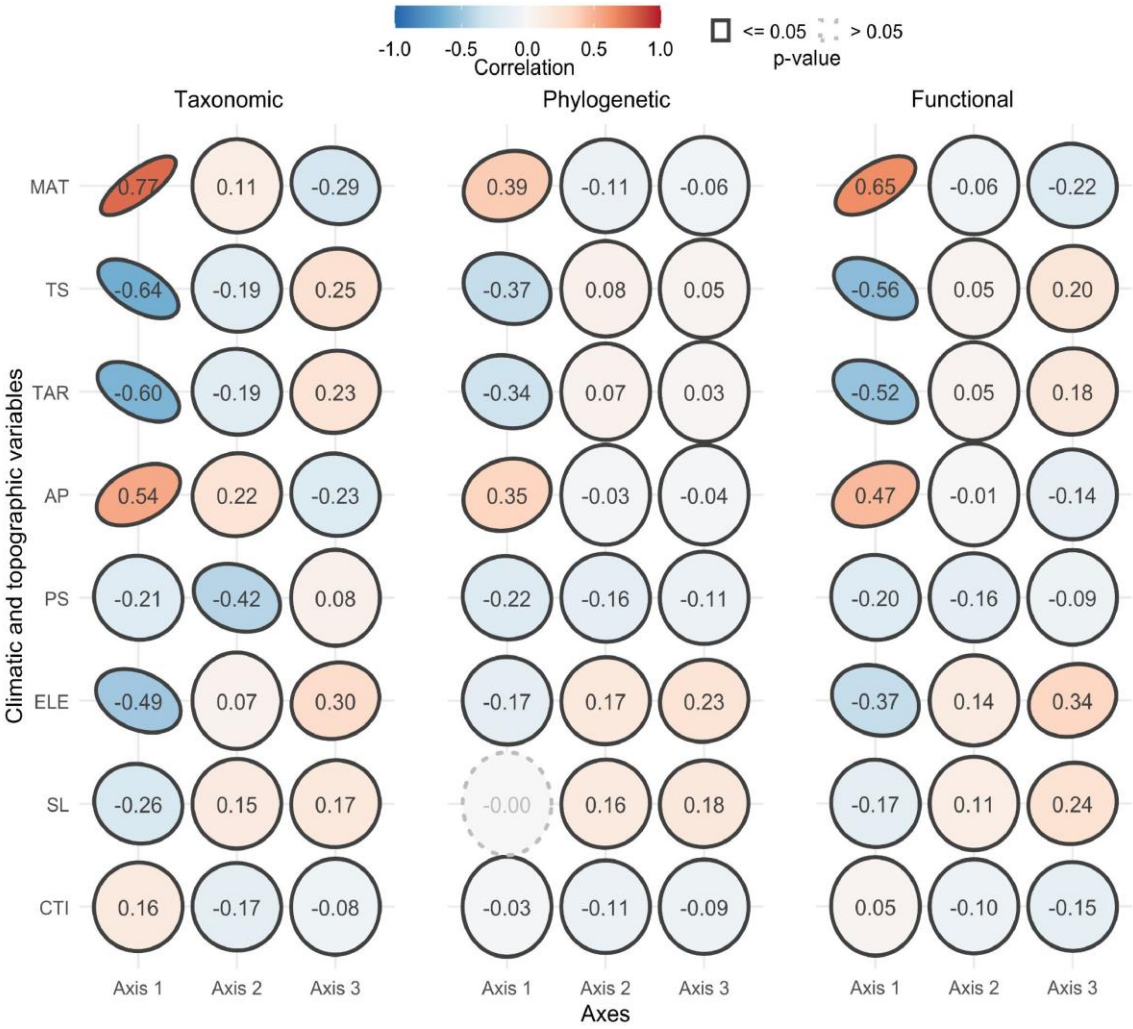


Fig. S9. Correlation between axes of Multi-Dimensional Scaling ordinations derived from taxonomic, phylogenetic, and functional beta diversity and climatic and topographic variables of forest communities. Acronyms represent the mean annual temperature (MAT), temperature seasonality (TS), temperature annual range (TAR), annual precipitation (AP), precipitation seasonality (PS), elevation (ELE), slope (SL), and compound topographic index (CTI).

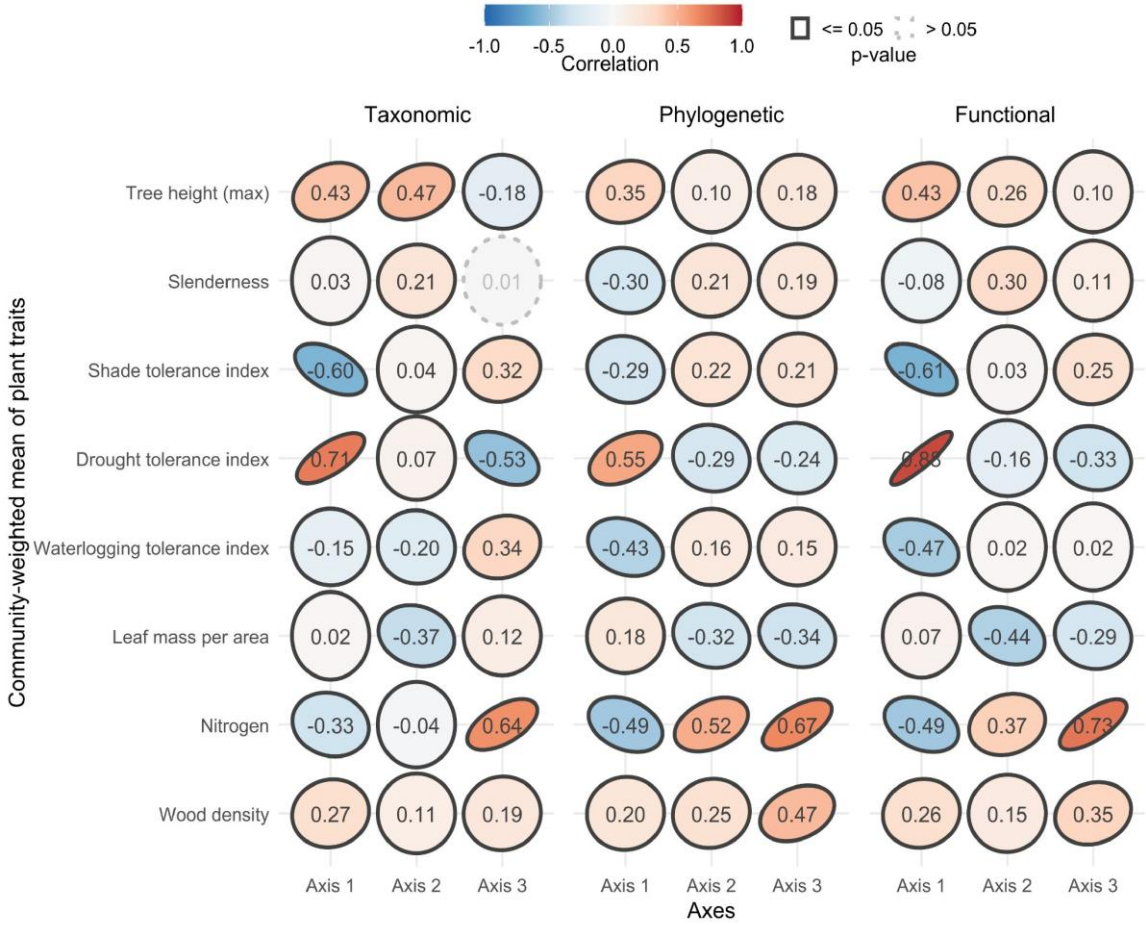


Fig. S10. Correlation between axes of Multi-Dimensional Scaling ordinations derived from taxonomic, phylogenetic, and functional beta diversity and community-weighted mean of plant traits of forest communities.

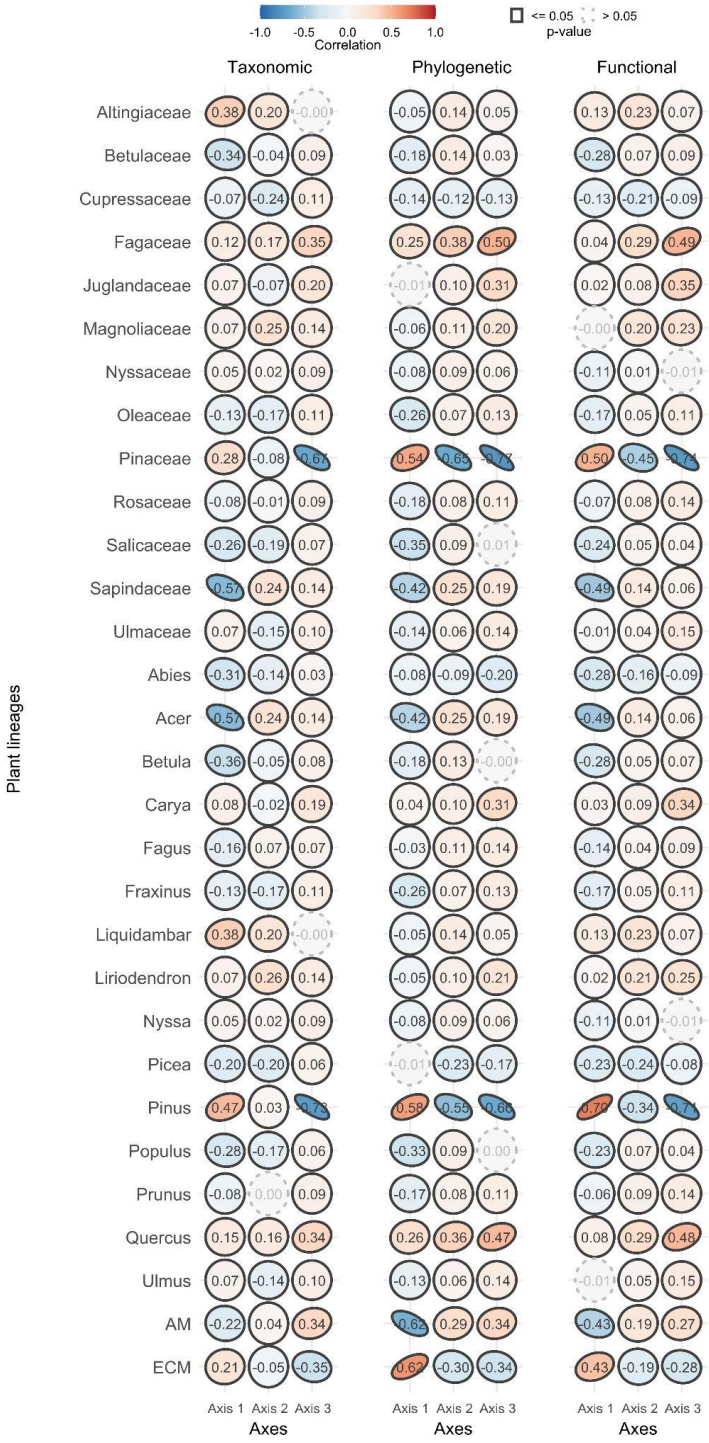


Fig. S11. Correlation between axes of Multi-Dimensional Scaling ordinations derived from taxonomic, phylogenetic, and functional beta diversity and the relative abundance of plant lineages within forest communities.

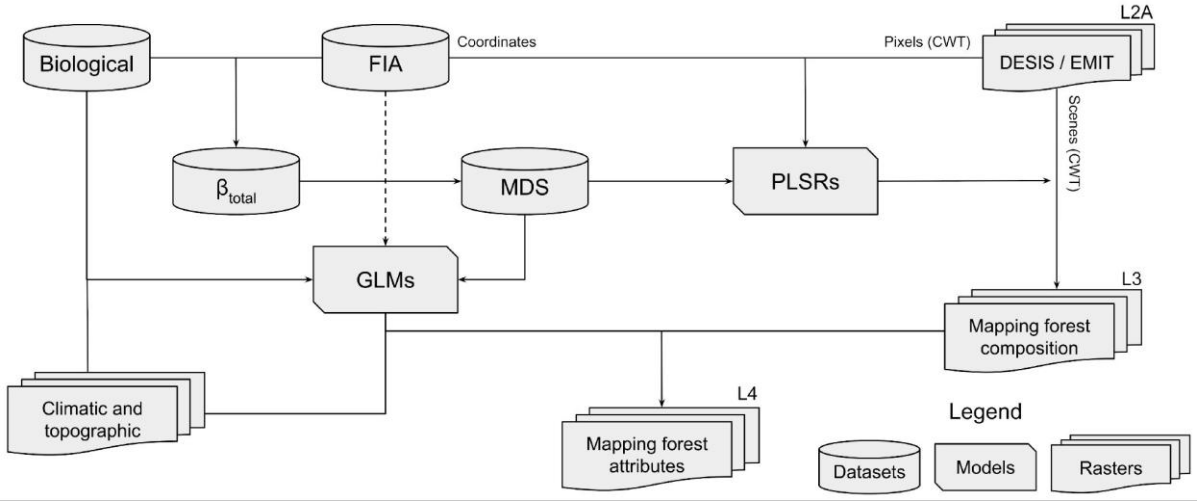


Fig. S12. Schematic representation of the workflow used to map multiple dimensions of forest diversity.

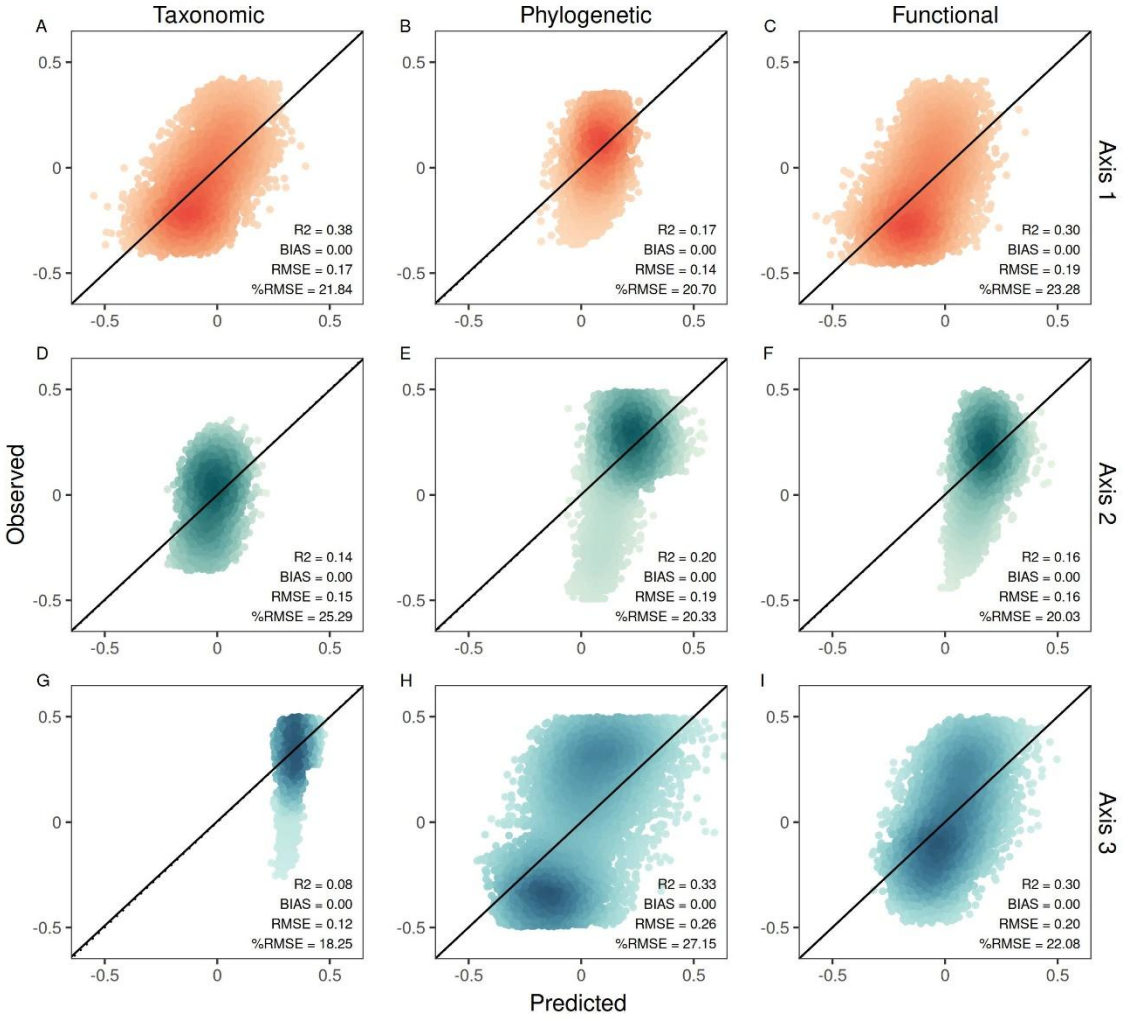


Fig. S13. Performance of training datasets of PLSR models that predict ordination axes of multiple dimensions of tree diversity using spaceborne observations from DESIS. Values represent the mean and standard deviation of 100 models.

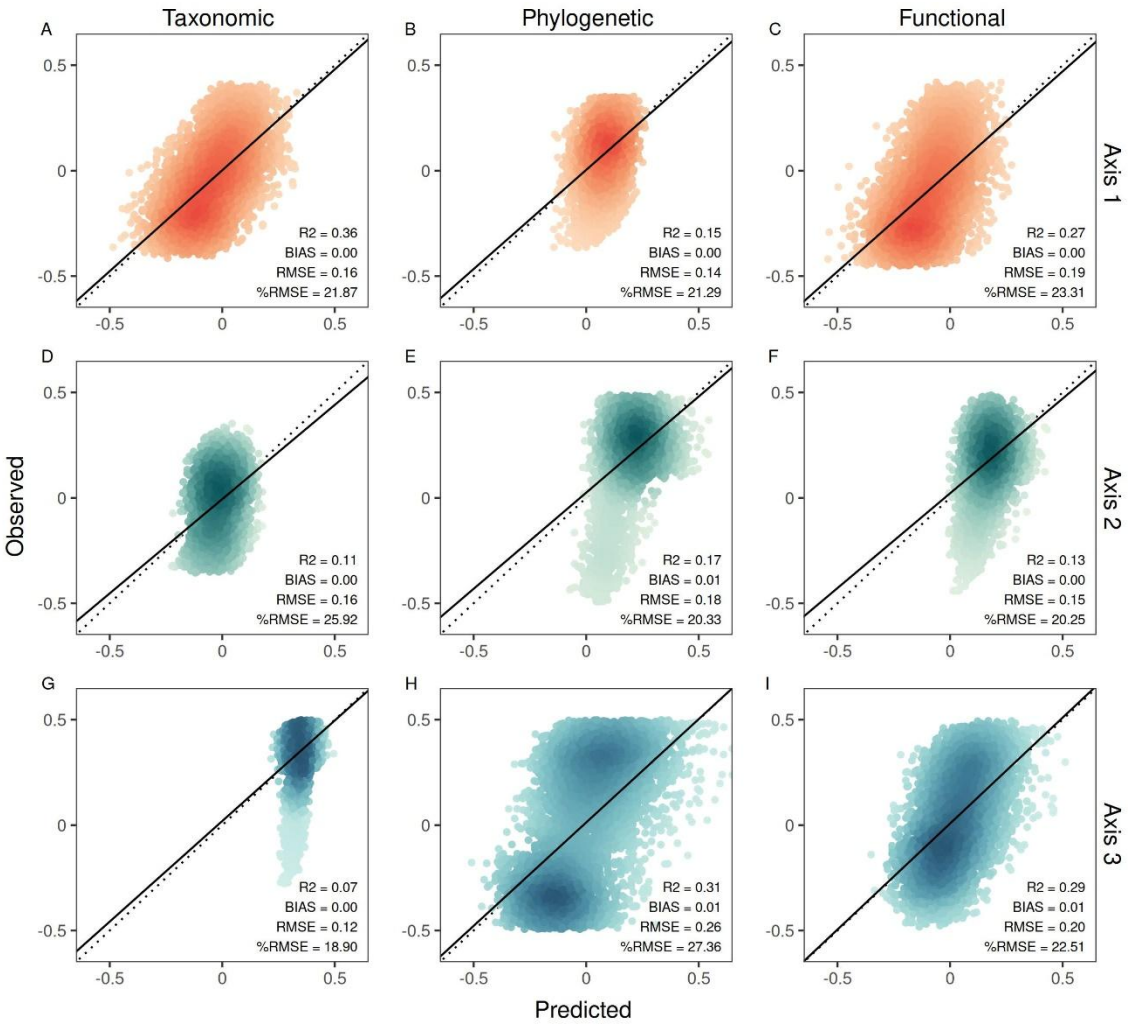


Fig. S14. Performance of testing datasets of PLSR models that predict ordination axes of multiple dimensions of tree diversity using spaceborne observations from DESIS. Values represent the mean and standard deviation of 100 models.

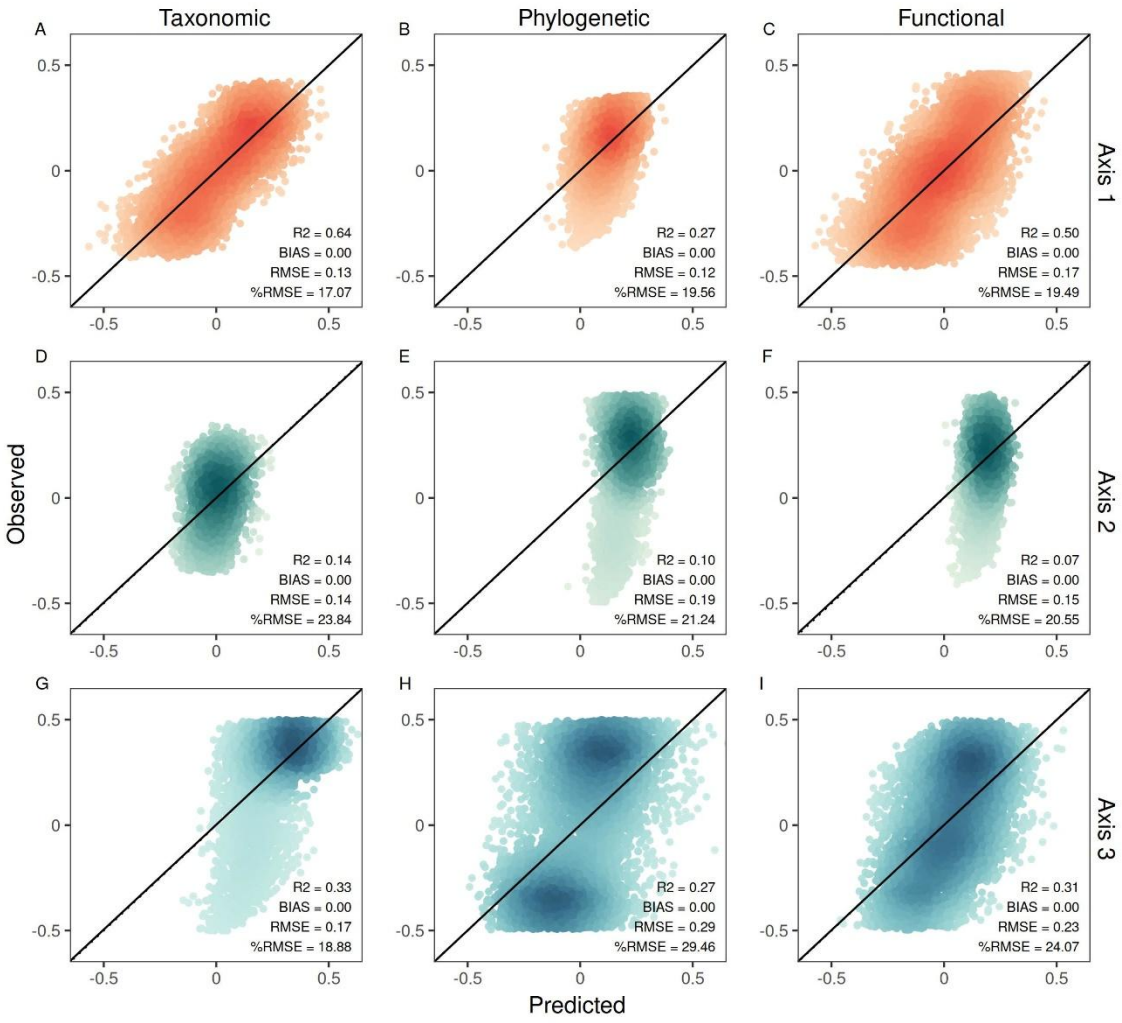


Fig. S15. Performance of training datasets of PLSR models that predict ordination axes of multiple dimensions of tree diversity using spaceborne observations from EMIT. Values represent the mean and standard deviation of 100 models.

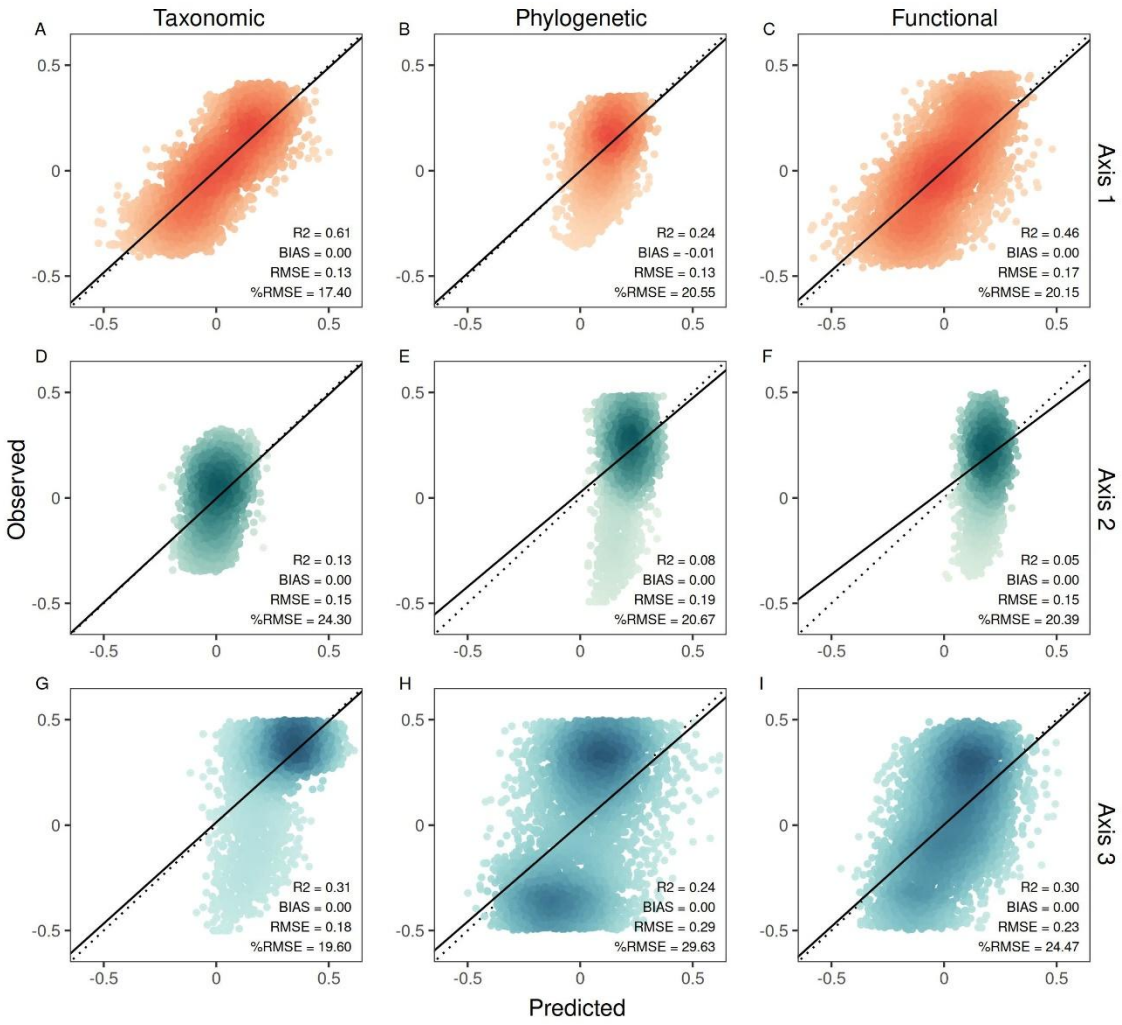


Fig. S16. Performance of testing datasets of PLSR models that predict ordination axes of multiple dimensions of tree diversity using spaceborne observations from EMIT. Values represent the mean and standard deviation of 100 models.

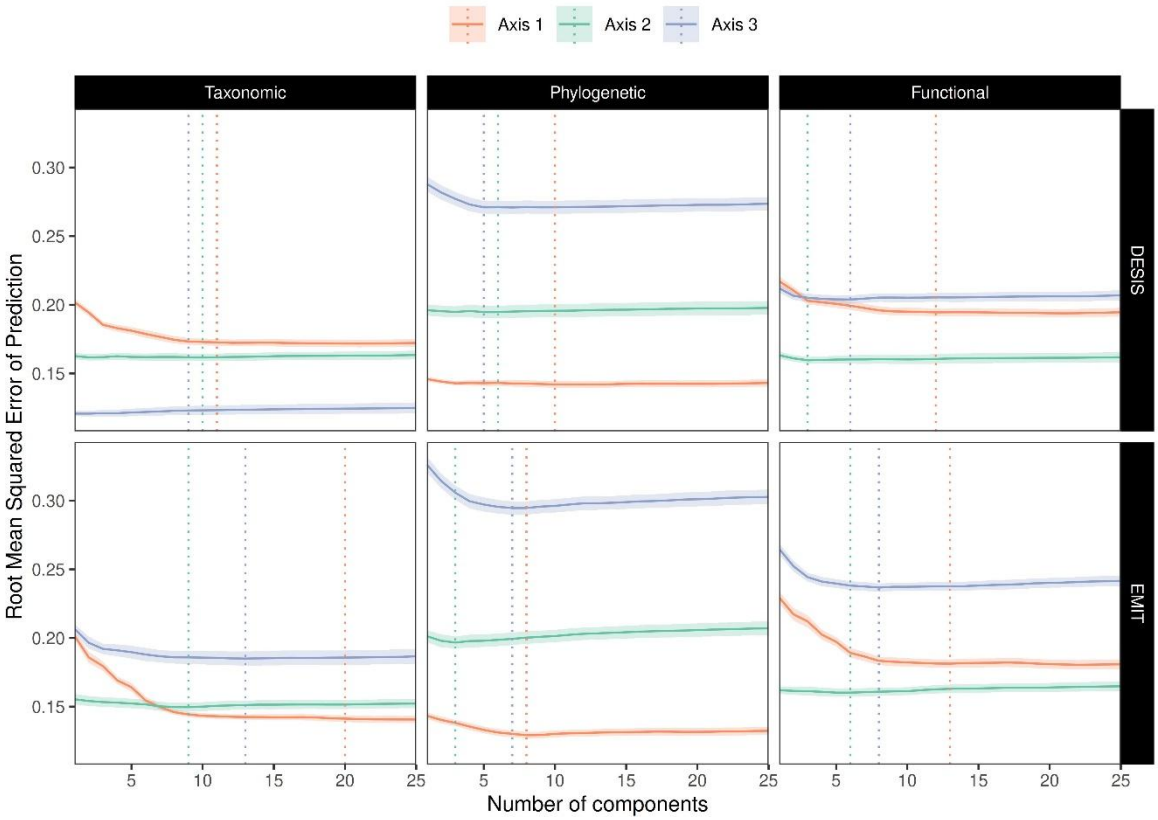


Fig. S17. Root Mean Squared Error of Prediction (RMSEP) from cross-validation derived from PLSR models that predict ordination axes of multiple dimensions of tree diversity using spaceborne observations from DESIS and EMIT. Each solid line represents the average of the 100 iterations, while the shade around each line is its standard deviation. Horizontal dotted lines represent the optimal number of components selected for each axis.

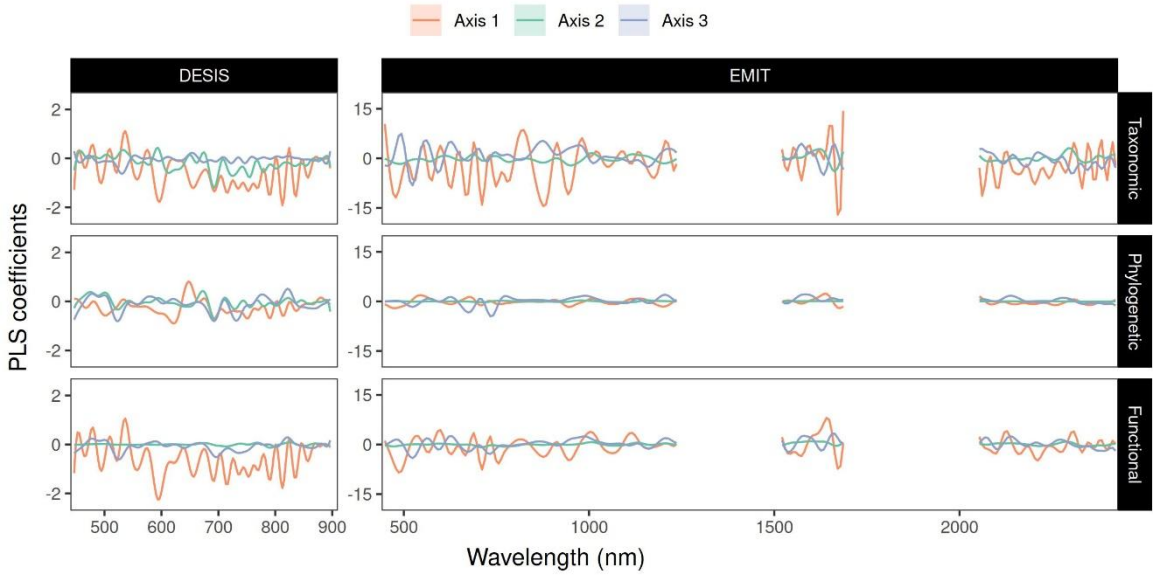


Fig. S18. PLSR coefficients of the models that predict ordination axes of multiple dimensions of tree diversity using wavelet-transformed spaceborne reflectance from DESIS and EMIT. Each line represents an average of 100 iterations, while the shade around each line represents the standard deviation.

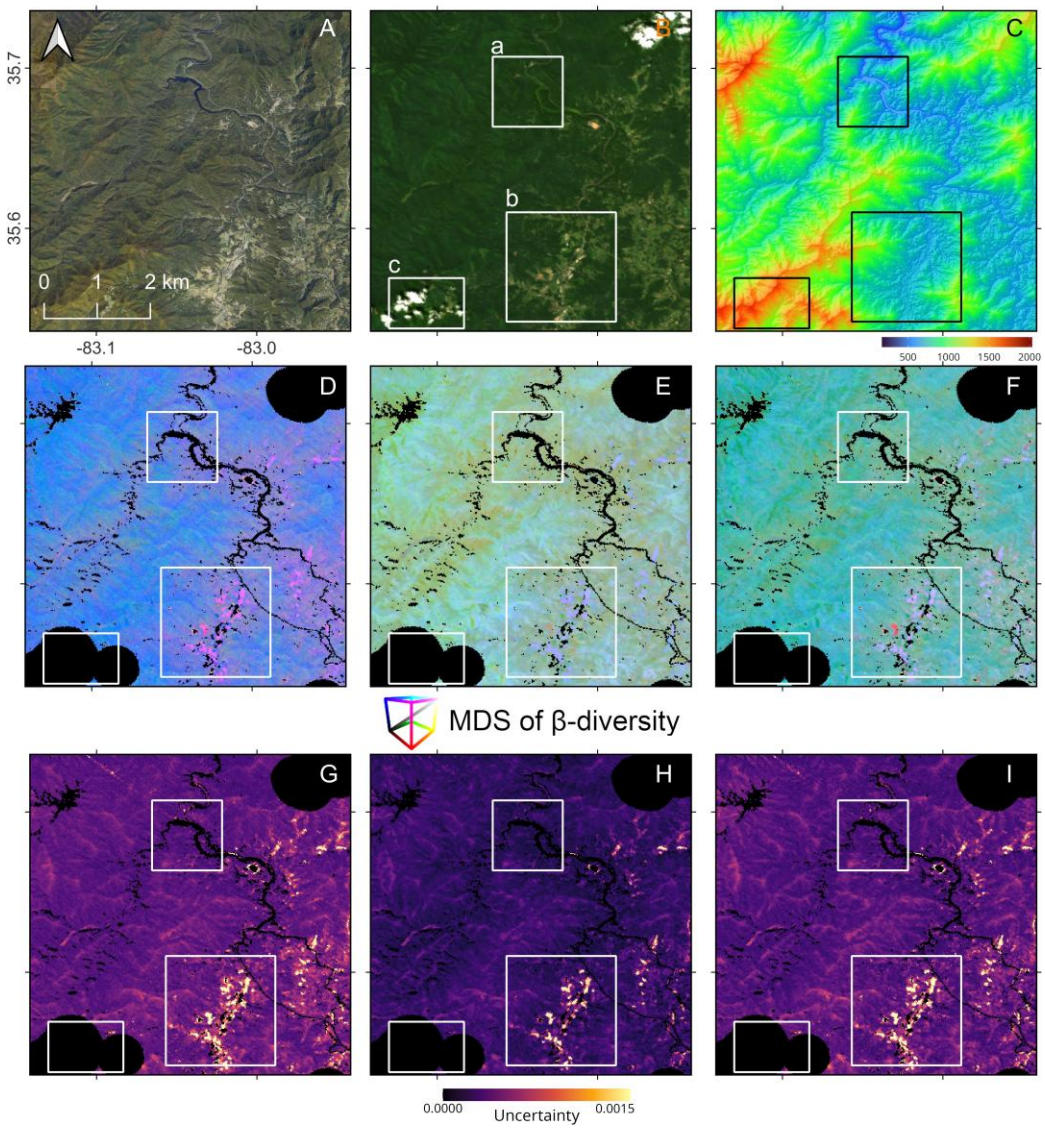


Fig. S19. Example of pixel masking using non-forest thresholds (a) and product flags (c), and the effect of spectral mixing on estimated uncertainty (d) in an EMIT L2 scene from the southern Appalachian Mountains. Panel A shows a high-resolution image from the National Agriculture Imagery Program for comparison with B. Panels B and C present a true-color image (R: λ663, G: λ551, B: λ425) and the elevation model provided by NASA alongside the EMIT imagery. Panels D, E, and F show the mapped MDS β-diversity axes based on taxonomic, phylogenetic, and functional dimensions of diversity, while panels G, H, and I show their associated uncertainty. The square area 'a' illustrates how NDVI and NIR thresholds are used to remove non-forest pixels when applying the models to scenes. The rectangle area 'c' demonstrates how the aggregated EMIT L2 mask flag is applied to remove potential cloud and cloud-shadow contamination from the scene and derived maps. The square area 'd' highlights cases where non-forest thresholds do not reliably filter pixels because of their spectral similarity to forest cover or spectral mixing, resulting in areas with high prediction uncertainty. These uncertainty maps can be used by users to constrain pixel selection depending on their objectives.

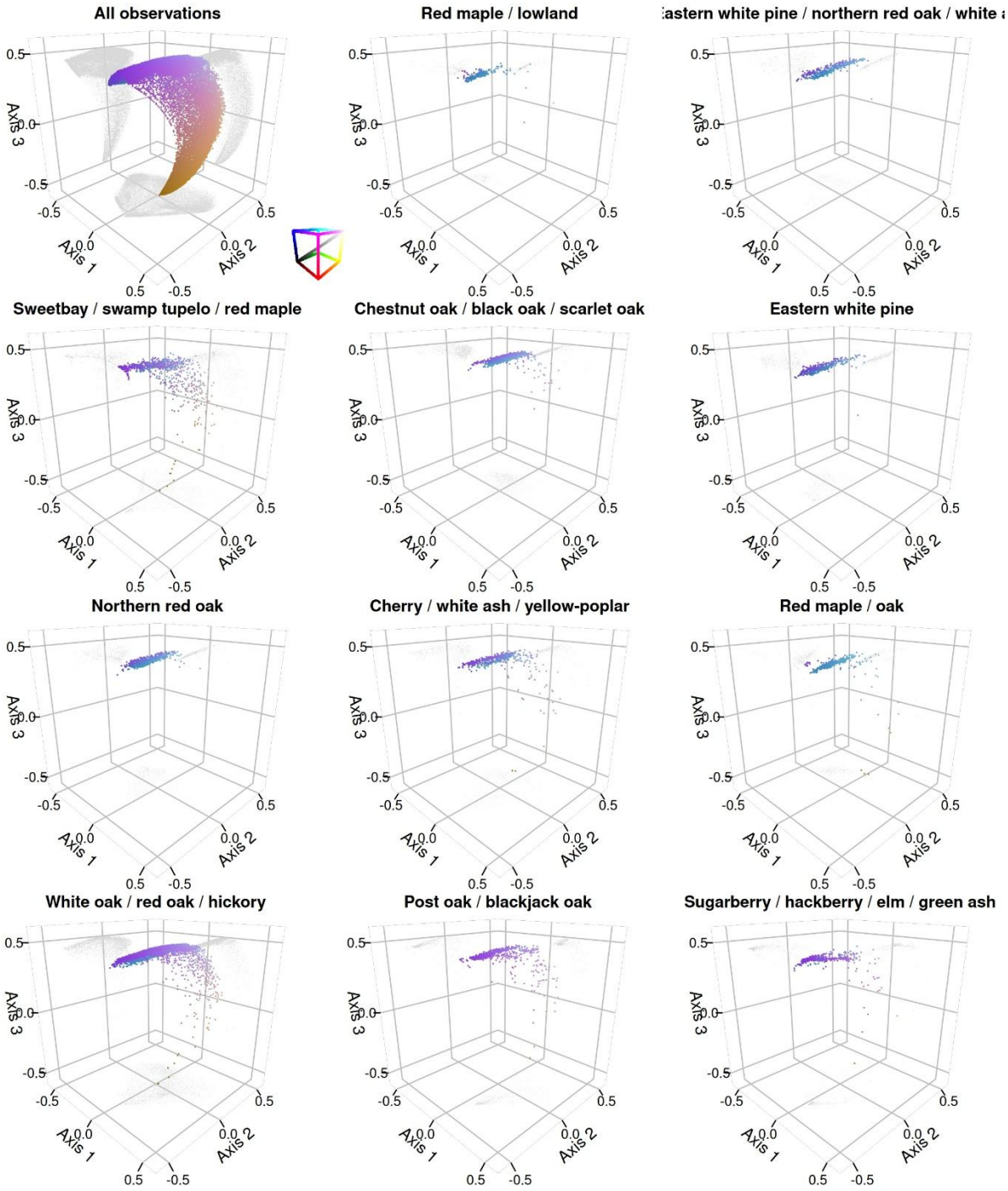
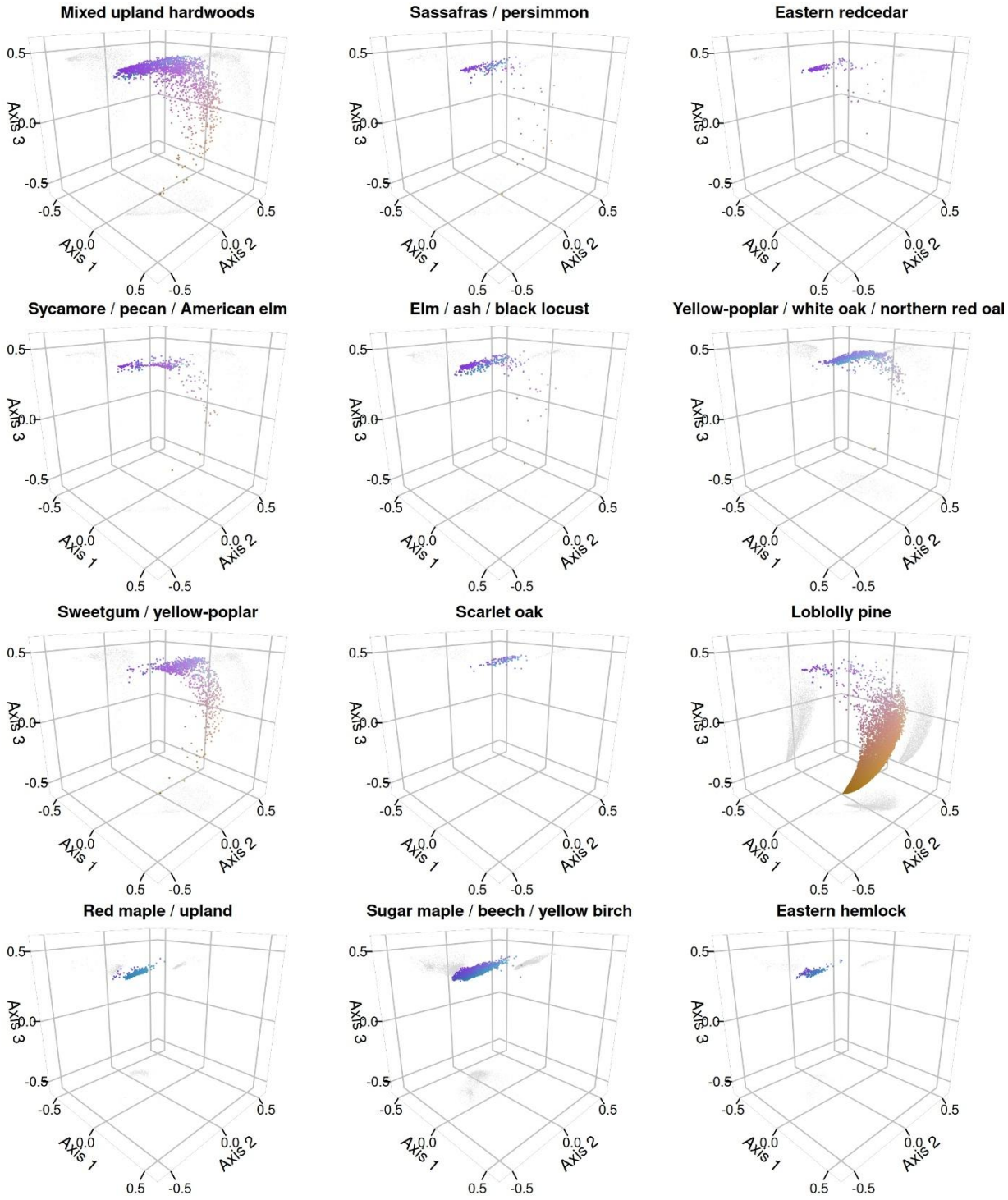
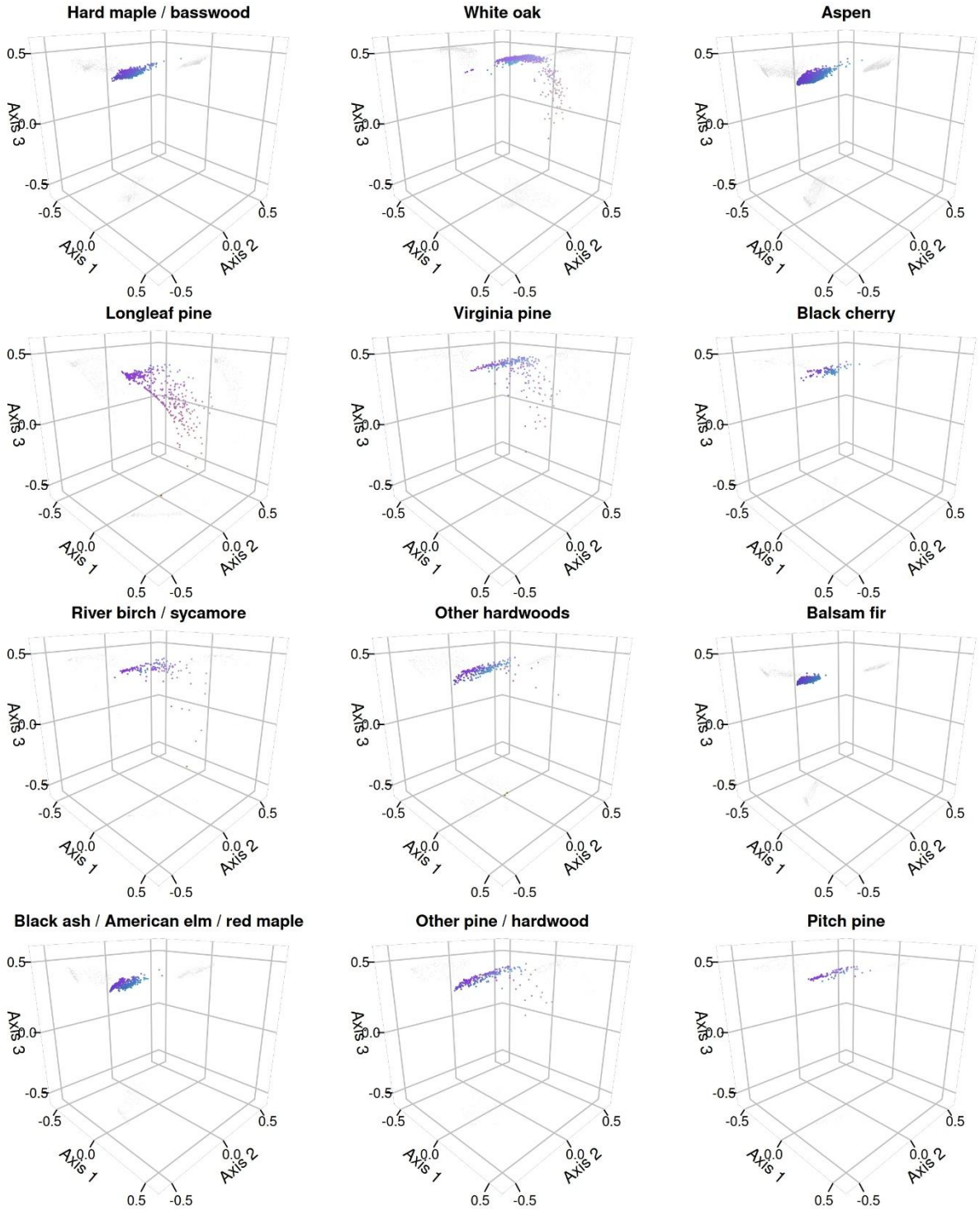
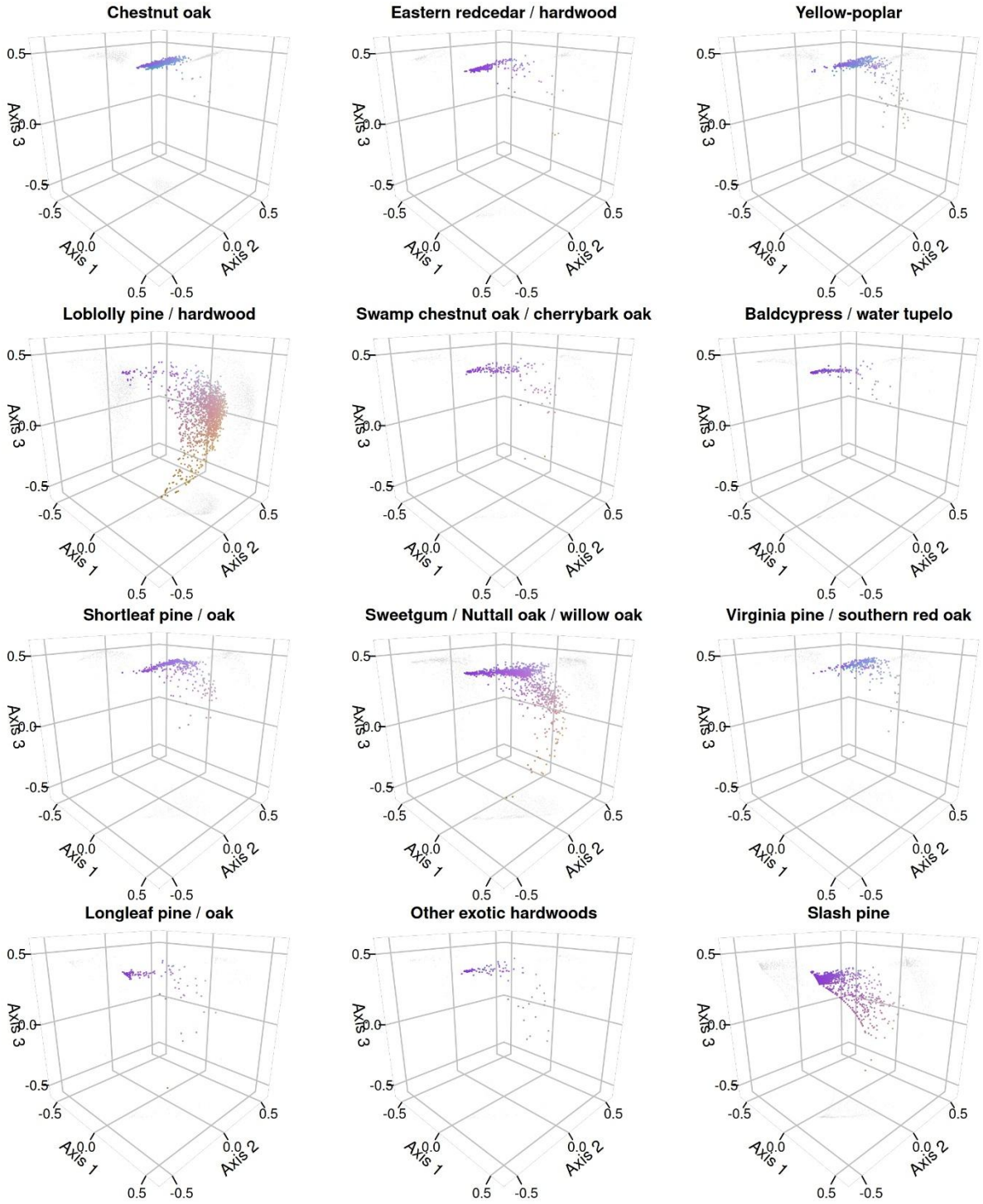
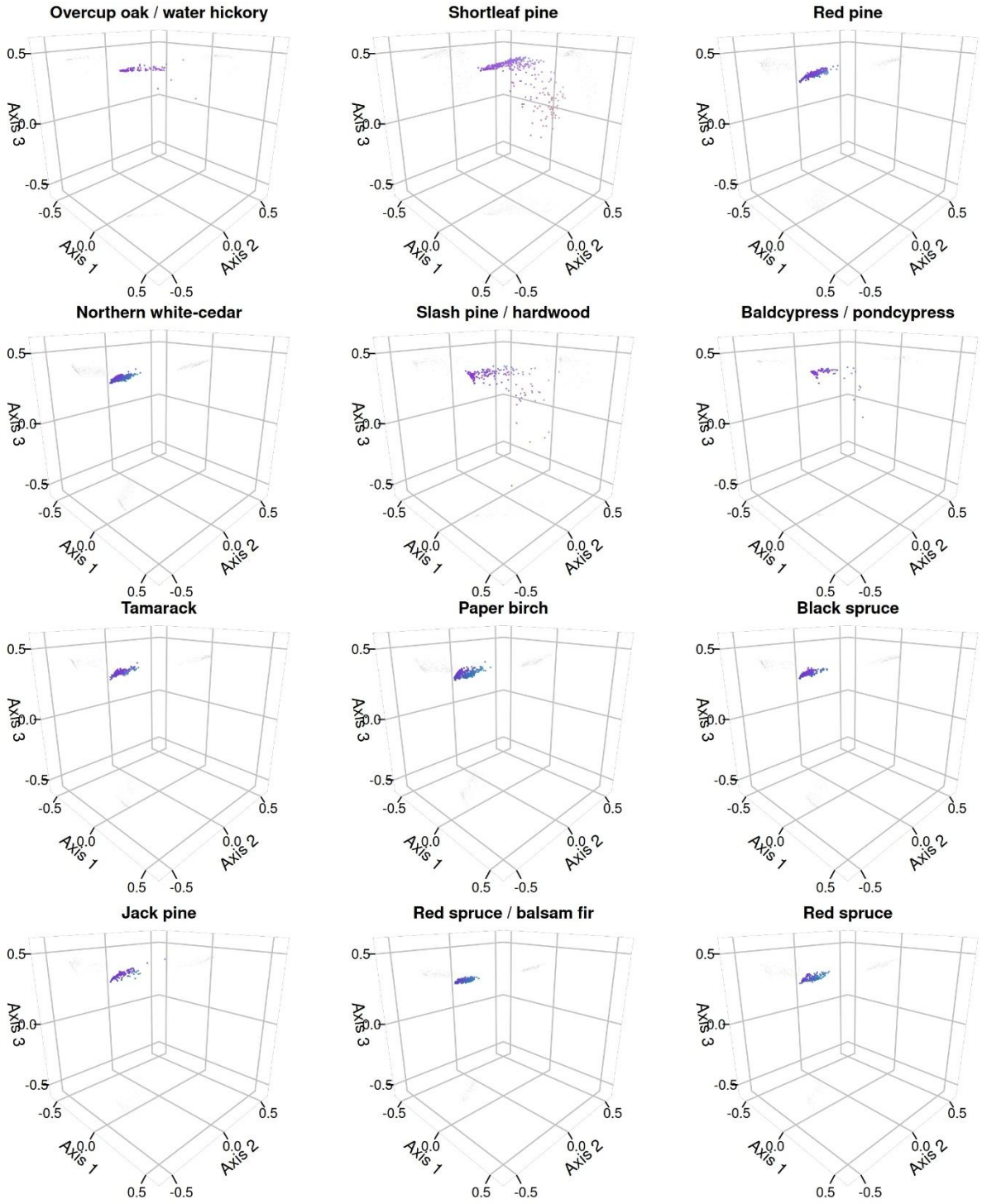


Fig. S20. Multi-dimensional scaling (MDS) used to ordinate taxonomic β-diversity. The first panel shows all forest communities, as shown in Fig. 3. The other panels show forest communities labeled with a FIA forest type. (Continues on the following pages).


Mixed upland hardwoods
Sassafras / persimmon
Eastern redcedar
Sycamore / pecan / American elm
Elm / ash / black locust
Yellow-poplar / white oak / northern red oal
Sweetgum / yellow-poplar
Scarlet oak
Loblolly pine
Red maple / upland
Sugar maple / beech / yellow birch
Eastern hemlock
Axis 1
Axis 2
Axis 3
0.5
0.0
-0.5


Hard maple / basswood
White oak
Aspen
Longleaf pine
Virginia pine
Black cherry
River birch / sycamore
Other hardwoods
Balsam fir
Black ash / American elm / red maple
Other pine / hardwood
Pitch pine
Axis 1
Axis 2
Axis 3
0.5
0.0
-0.5


Chestnut oak
Eastern redcedar / hardwood
Yellow-poplar
Loblolly pine / hardwood
Swamp chestnut oak / cherrybark oak
Baldcypress / water tupelo
Shortleaf pine / oak
Sweetgum / Nuttall oak / willow oak
Virginia pine / southern red oak
Longleaf pine / oak
Other exotic hardwoods
Slash pine
Axis 1
Axis 2
Axis 3
0.5
0.0
-0.5


Overcup oak / water hickory
Shortleaf pine
Red pine
Northern white-cedar
Slash pine / hardwood
Baldcypress / pondcypress
Tamarack
Paper birch
Black spruce
Jack pine
Red spruce / balsam fir
Red spruce
Axis 1
Axis 2
Axis 3
0.5
0.0
-0.5

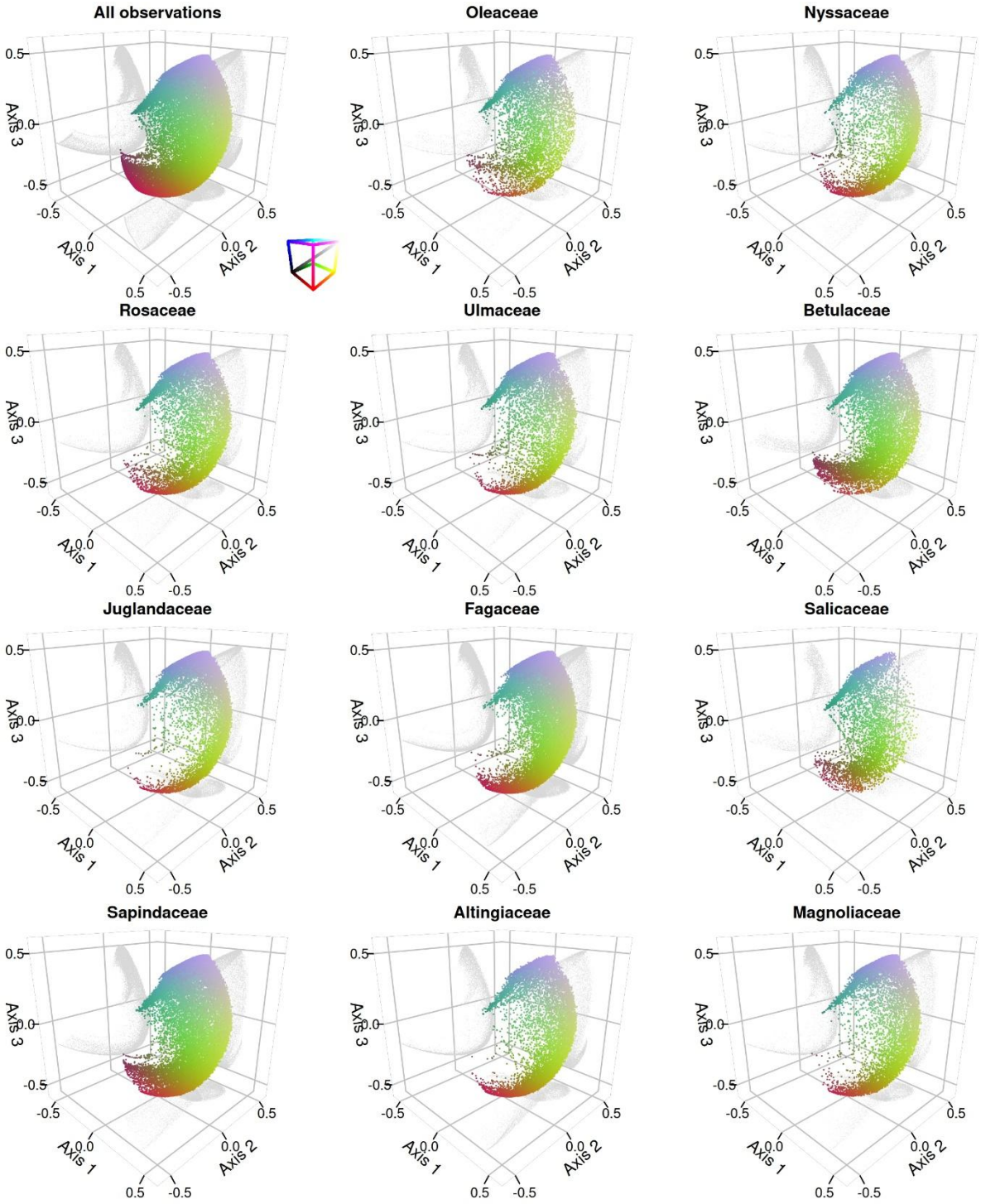
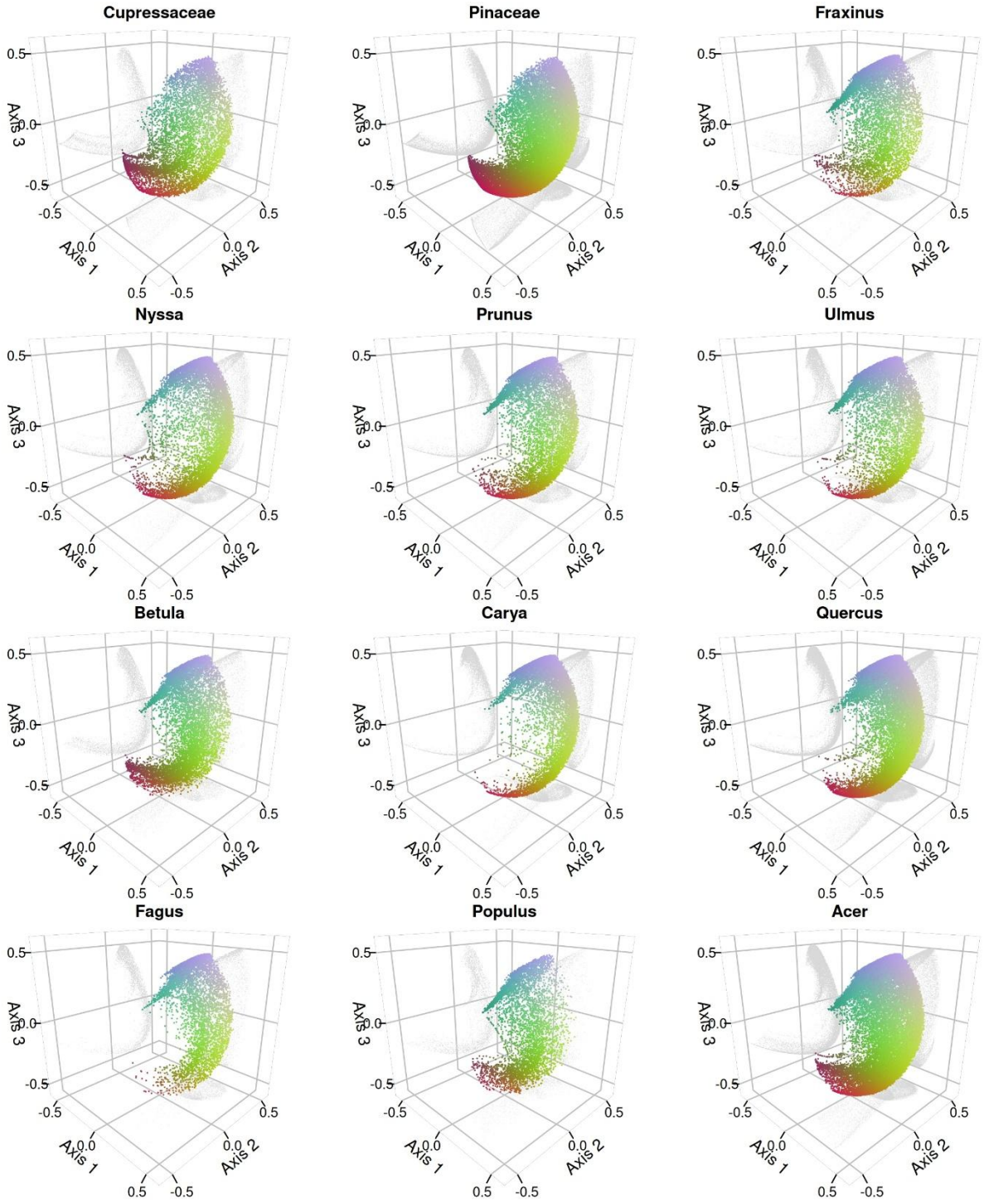
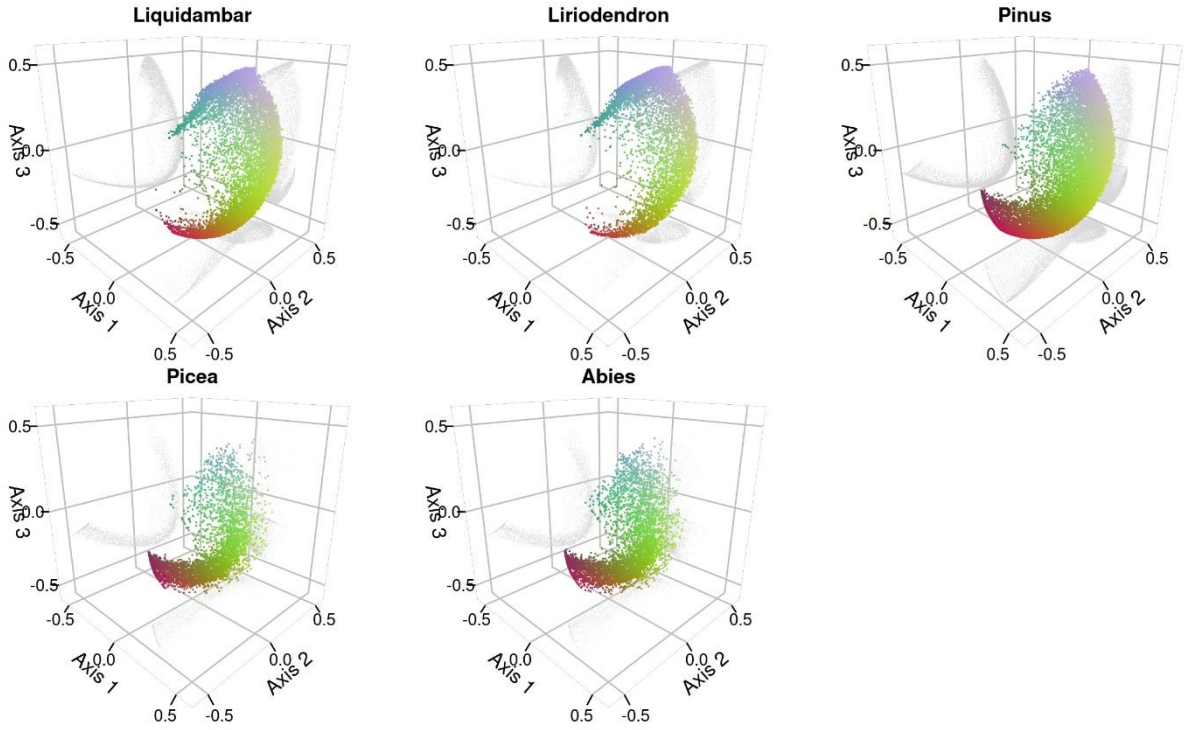


Fig. S21. Multi-dimensional scaling (MDS) used to ordinate phylogenetic β-diversity. The first panel shows all forest communities, as shown in Fig. 3. The other panels show forest communities with the occurrence of at least one tree from a given family or genera (Continues on the following pages).


Cupressaceae
Pinaceae
Fraxinus
Nyssa
Prunus
Ulmus
Betula
Carya
Quercus
Fagus
Populus
Acer
Axis 1
Axis 2
Axis 3
0.5
0.0
-0.5


Liquidambar
Liriodendron
Pinus
Picea
Abies
Axis 1
Axis 2
Axis 3
0.5
0.0
-0.5

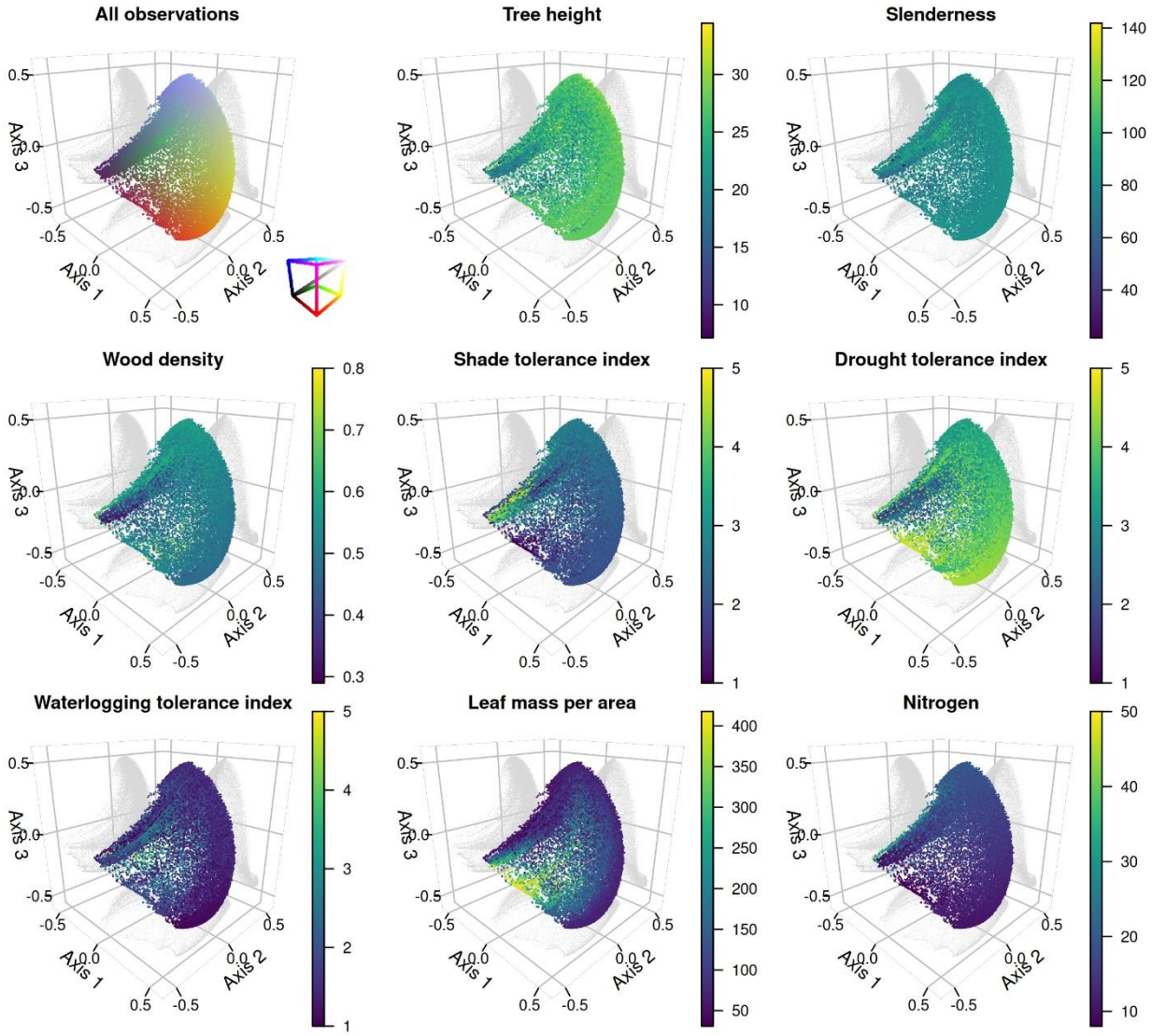


Fig. S22. Multi-dimensional scaling (MDS) used to ordinate functional β-diversity. The first panel shows all forest communities, as shown in Fig. 3. The other panels show forest communities with gradients of CWM of plant traits. The units are: tree height (m), slenderness (-), wood density (g/m^3), shade tolerance index (-), drought tolerance index (-), waterlogging tolerance index (-), leaf mass per area (g/m^2), and nitrogen (mg/g).

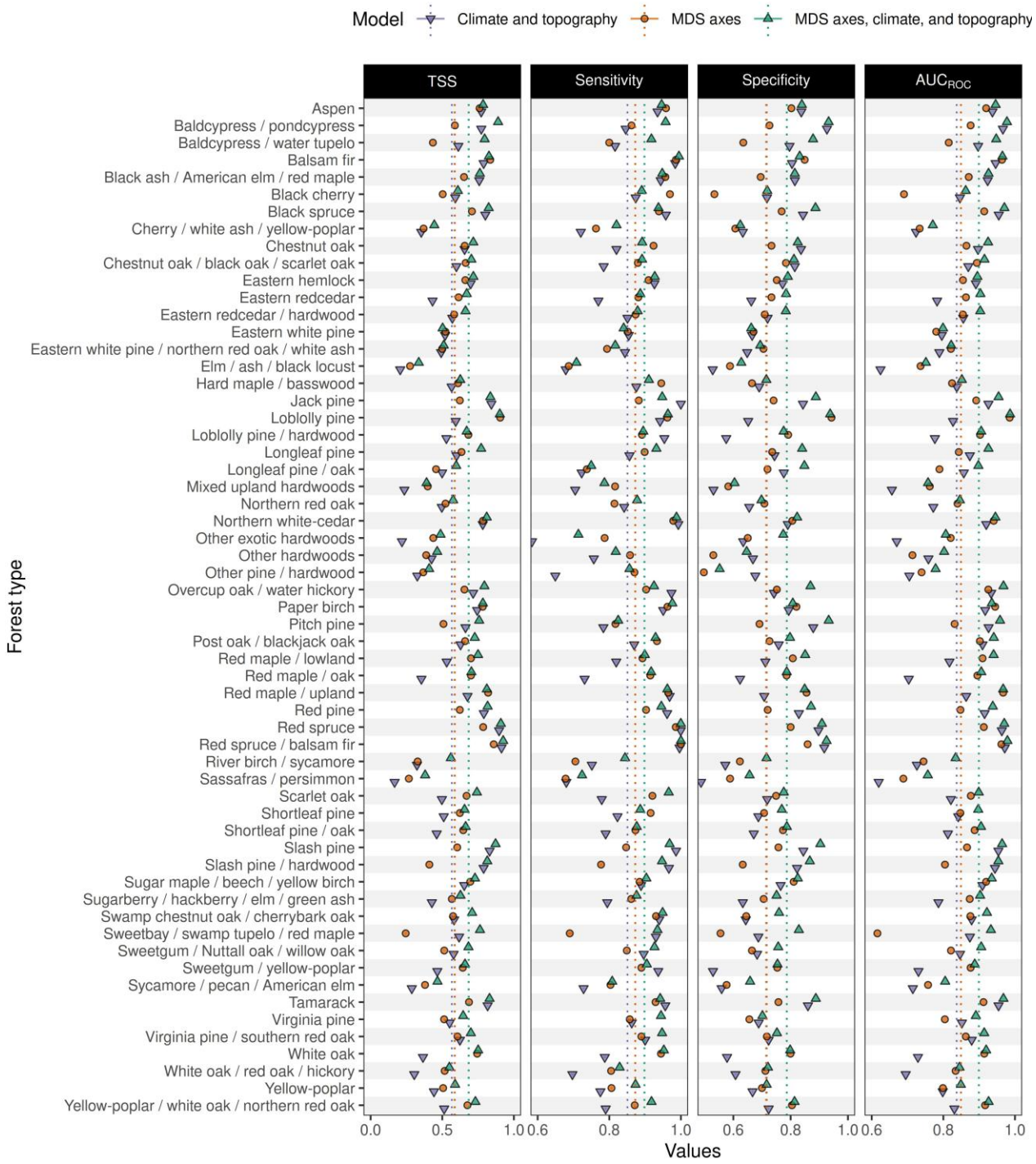


Fig. S23. Performance of binary generalized linear models used to predict forest type labels from FIA inventories. Three model configurations based on MDS ordination axes of beta diversity derived from taxonomic, climatic, and topographic data were used. Each point represents the mean estimate from 100 models.

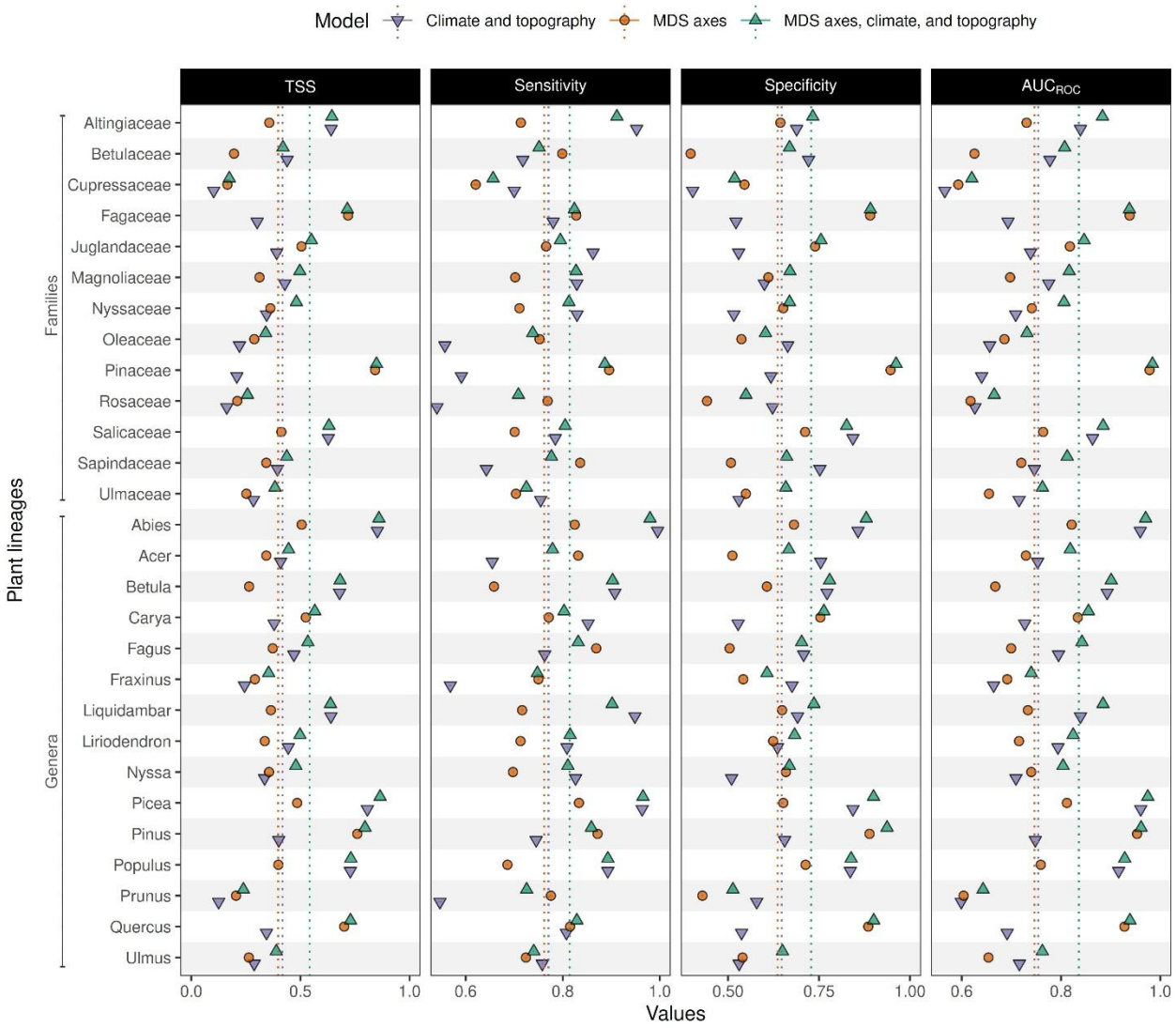


Fig. S24. Performance of binary generalized linear models used to predict the presence/absence of plant lineages within inventories from the FIA program. Three model configurations based on MDS ordination axes of beta diversity derived from taxonomic, climatic, and topographic data were used. Each point represents the mean estimate from 100 models.

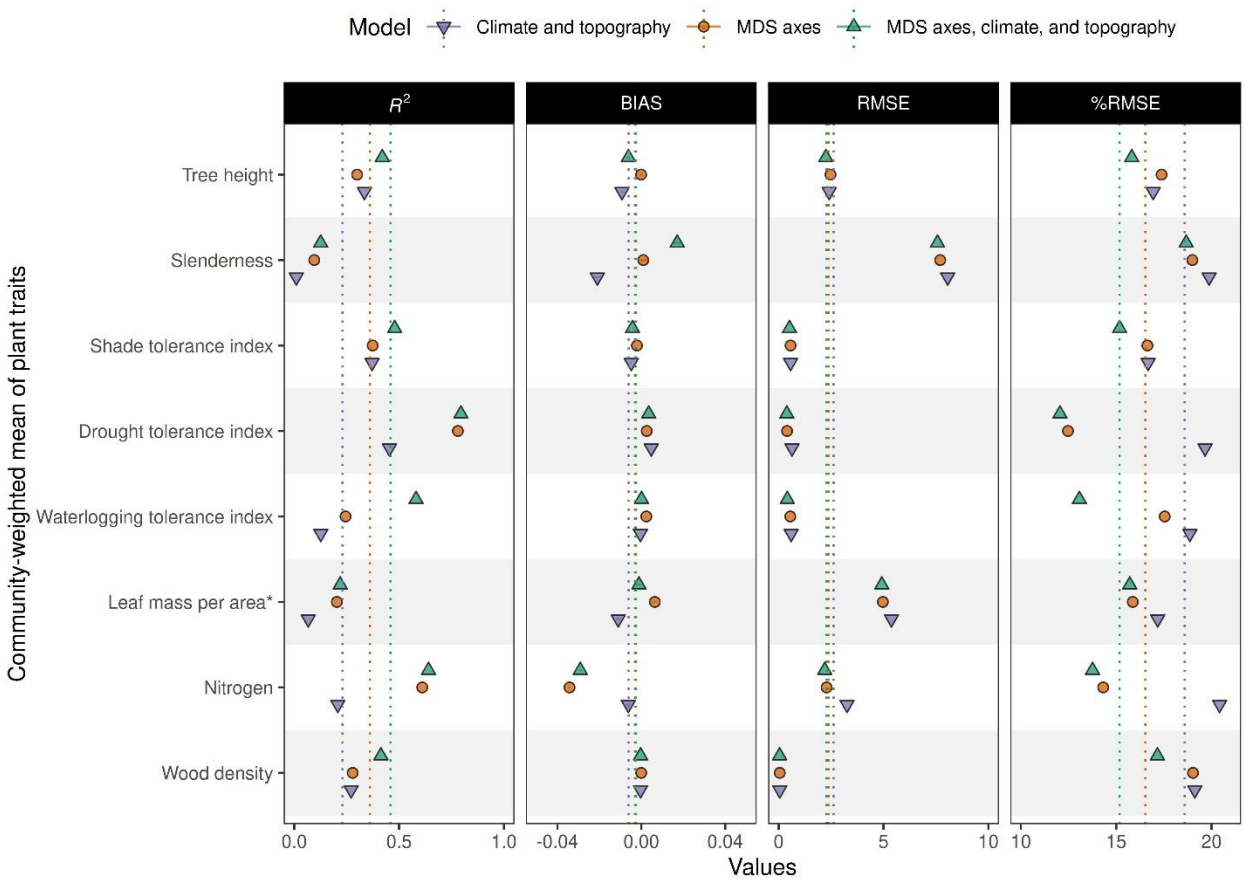


Fig. S25. Performance of generalized linear models used to predict community-weighted means of plant traits within inventories from the FIA program. Three model configurations based on MDS ordination axes of beta diversity derived from taxonomic, climatic, and topographic data were used. Each point represents the mean estimate from 100 models.

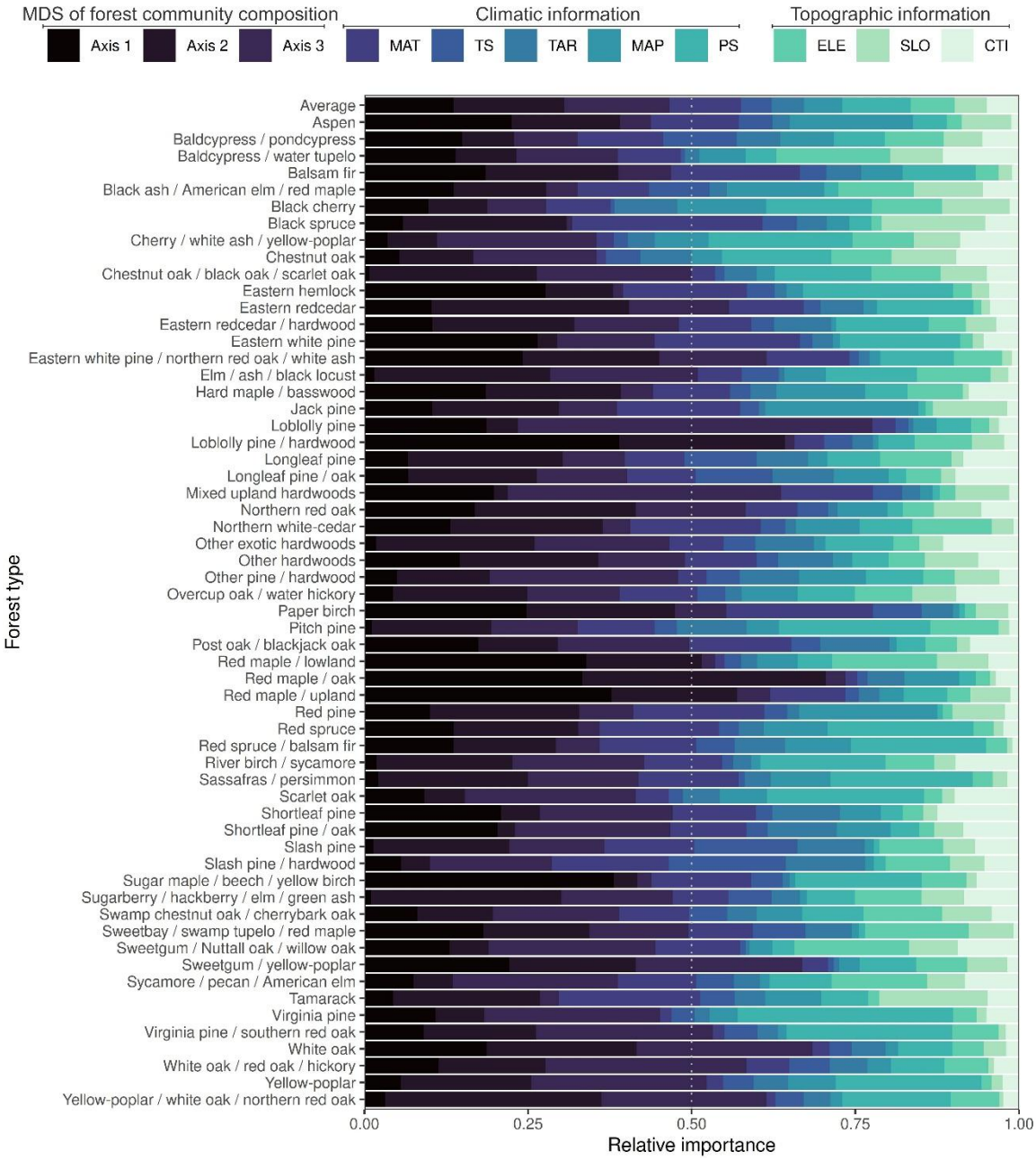


Fig. S26. Relative importance of variables used in binary generalized linear models to predict presence/absence of forest types within the Forest Inventory and Analysis program through axes of forest composition, climatic, and topographic information. Axes (i.e., Axis 1, Axis 2, and Axis 3) describe the potential forest composition from a MDS ordination of beta diversity based on taxonomic information. Climatic variables are described by the mean annual temperature (MAT), the temperature seasonality (TS), temperature annual range (TAR), annual precipitation (AP), and precipitation seasonality (PS). Topographic information is described by elevation (ELE), slope (SL), and compound topographic index (CTI).

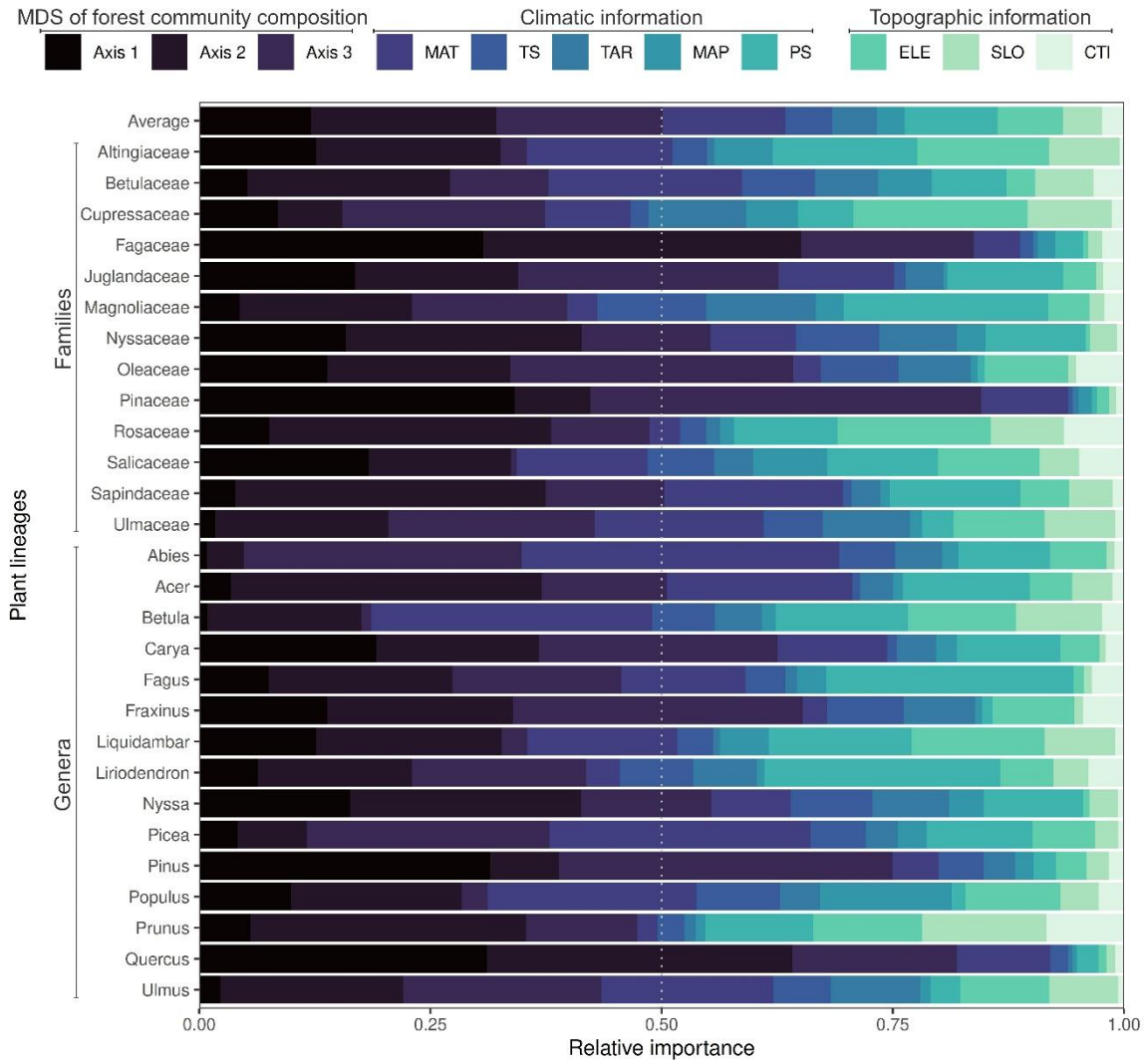


Fig. S27. Relative importance of variables used in binary generalized linear models to predict presence/absence of plant lineages within inventories of the Forest Inventory and Analysis program through axes of forest composition, climatic, and topographic information. Axes (i.e., Axis 1, Axis 2, and Axis 3) describe the potential forest composition from a MDS ordination of beta diversity based on phylogenetic information. Climatic variables are described by the mean annual temperature (MAT), the temperature seasonality (TS), temperature annual range (TAR), annual precipitation (AP), and precipitation seasonality (PS). Topographic information is described by elevation (ELE), slope (SL), and compound topographic index (CTI).

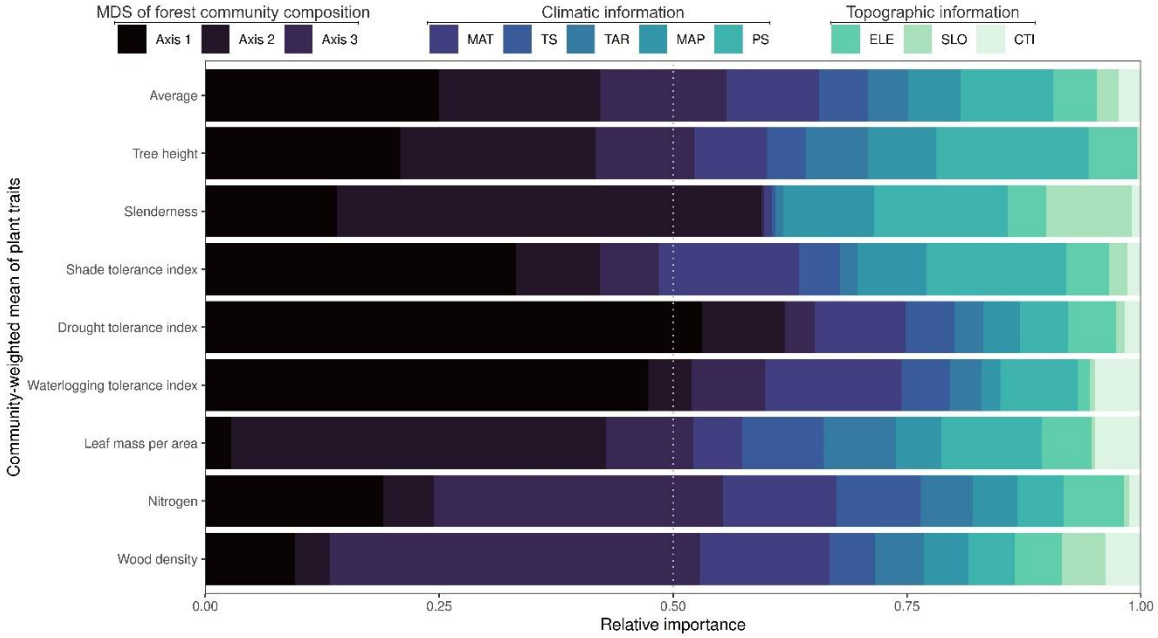


Fig. S28. Relative importance of variables used in generalized linear models to predict the community-weighted mean of plant traits within inventories from the Forest Inventory and Analysis program through axes of forest composition, climatic, and topographic information. Axes (i.e., Axis 1, Axis 2, and Axis 3) describe the potential forest composition from a MDS ordination of beta diversity based on functional information. Climatic variables are described by the mean annual temperature (MAT), the temperature seasonality (TS), temperature annual range (TAR), annual precipitation (AP), and precipitation seasonality (PS). Topographic information is described by elevation (ELE), slope (SL), and compound topographic index (CTI).

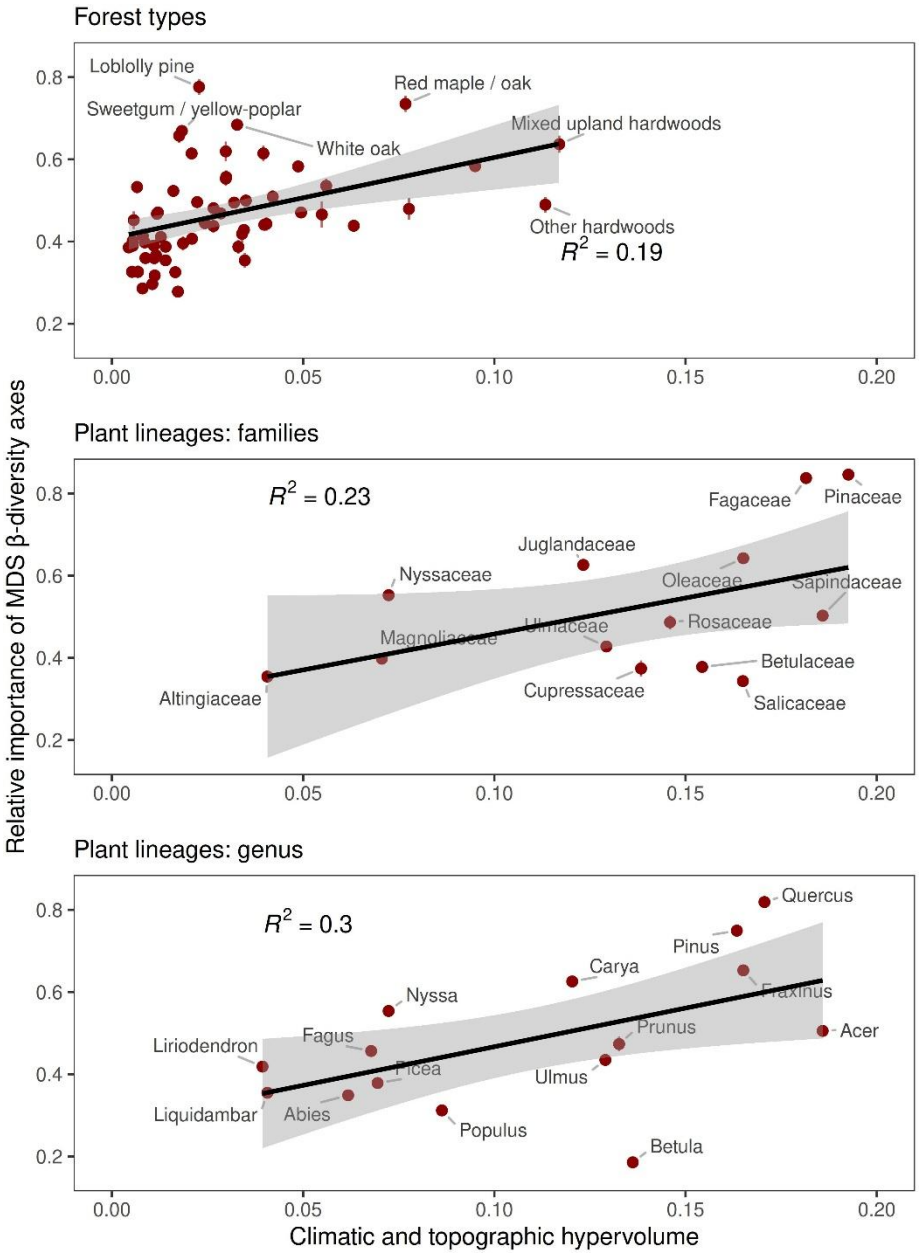


Fig. S29. Relationship between climatic and topographic hypervolume and the relative importance of MDS β-diversity axes for predicting forest attributes. Panels correspond to forest types (top), plant lineages at the family level (middle), and plant lineages at the genus level (bottom). Each point represents an individual attribute modelled, with the x-axis indicating its hypervolume and the y-axis indicating the relative importance of MDS axes in the corresponding model. Solid lines represent linear regressions, and shaded areas denote 95% confidence intervals. Reported R^2 values indicate the strength of the relationship in each panel.

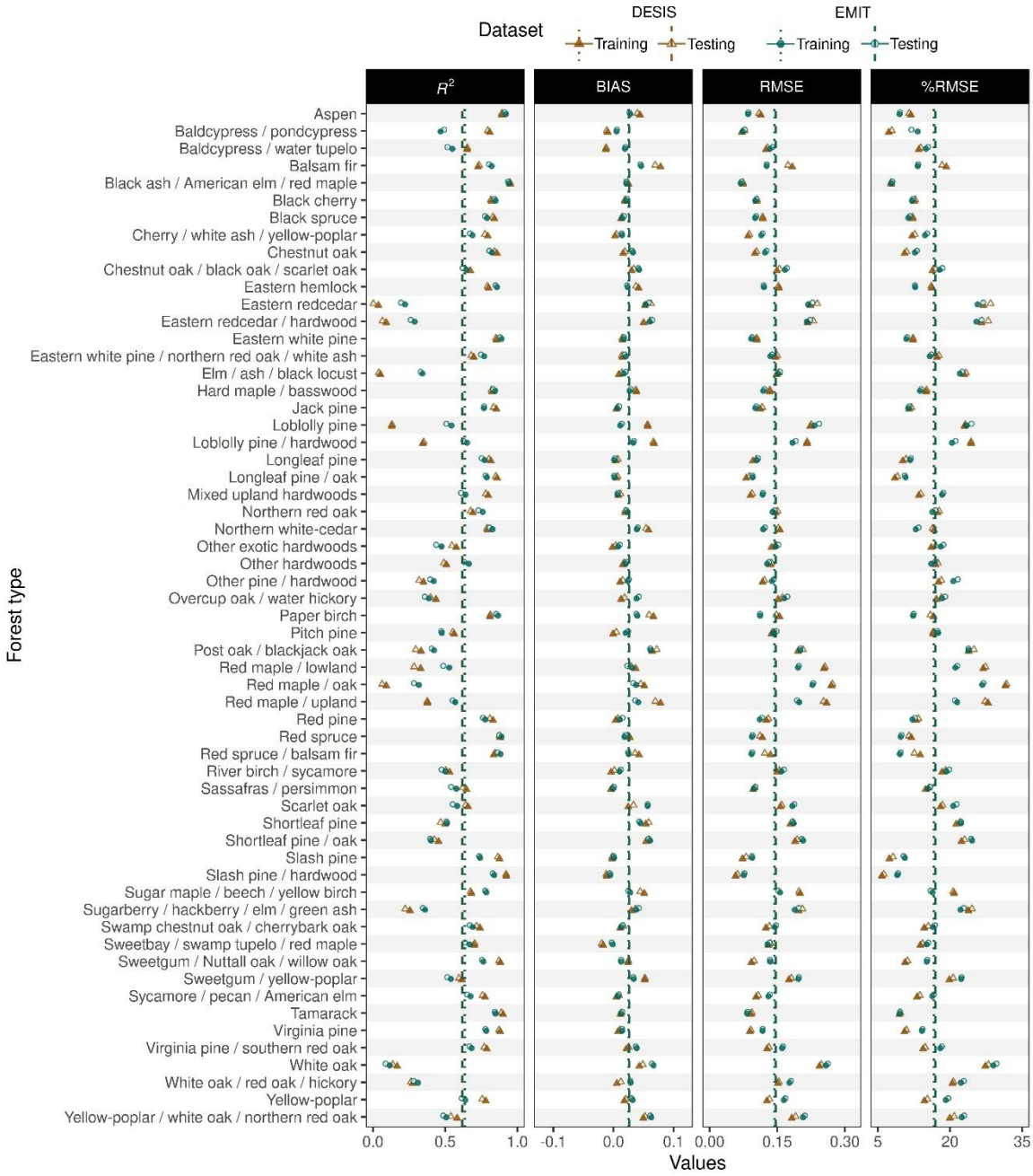


Fig. S30. Performance to predict the probability of forest types of inventories within the Forest Inventory and Analysis program using climatic and topographic information as well as predicted MDS axes of dimensions of beta diversity based on spaceborne observations of DESIS or EMIT. Each point represents the mean estimate of 100 models.

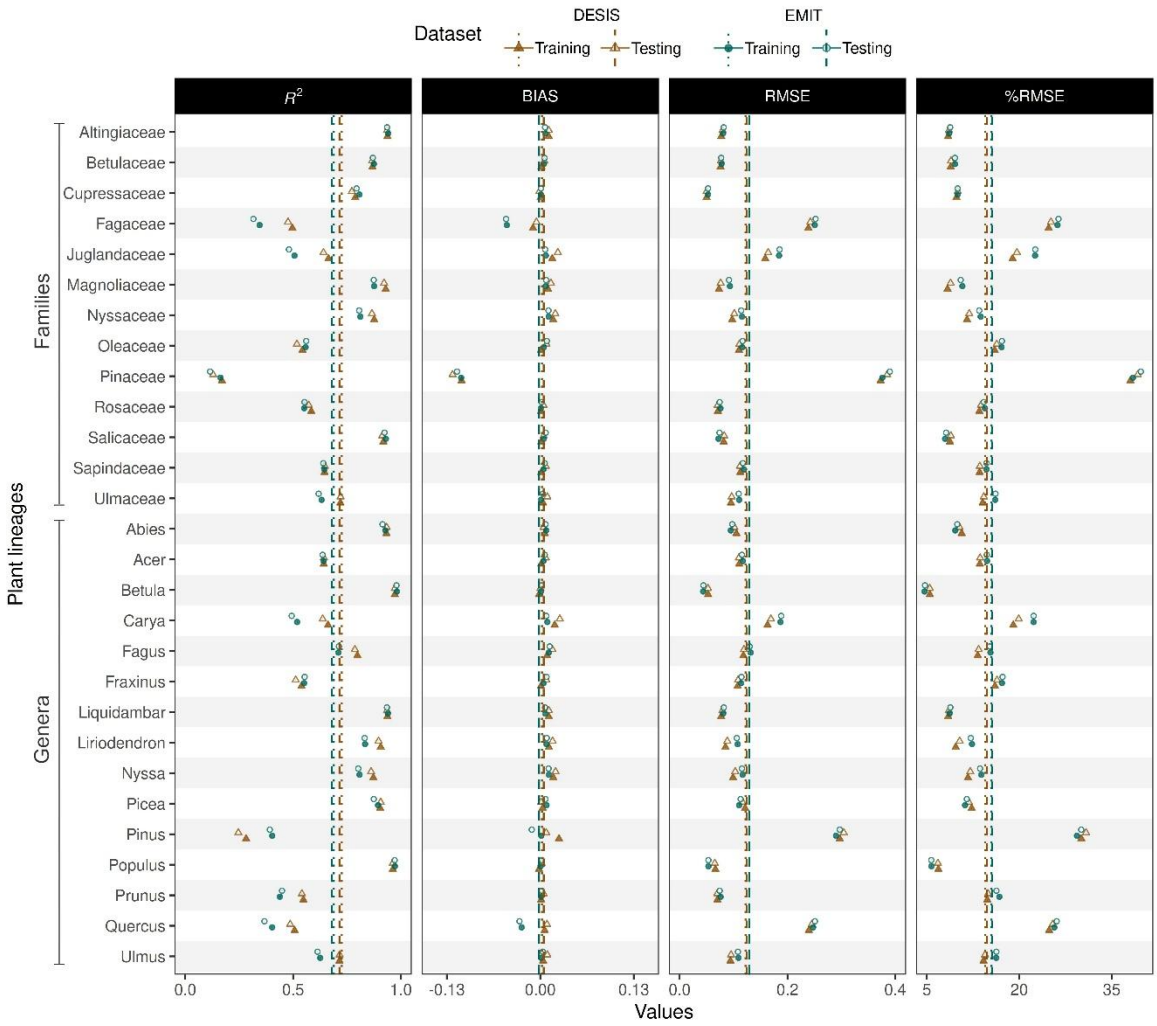


Fig. S31. Performance to predict the probability of presence of plant lineages within inventories of the Forest Inventory and Analysis program using climatic and topographic information as well as predicted MDS axes of dimensions of beta diversity based on spaceborne observations of DESIS or EMIT. Each point represents the mean estimate of 100 models.

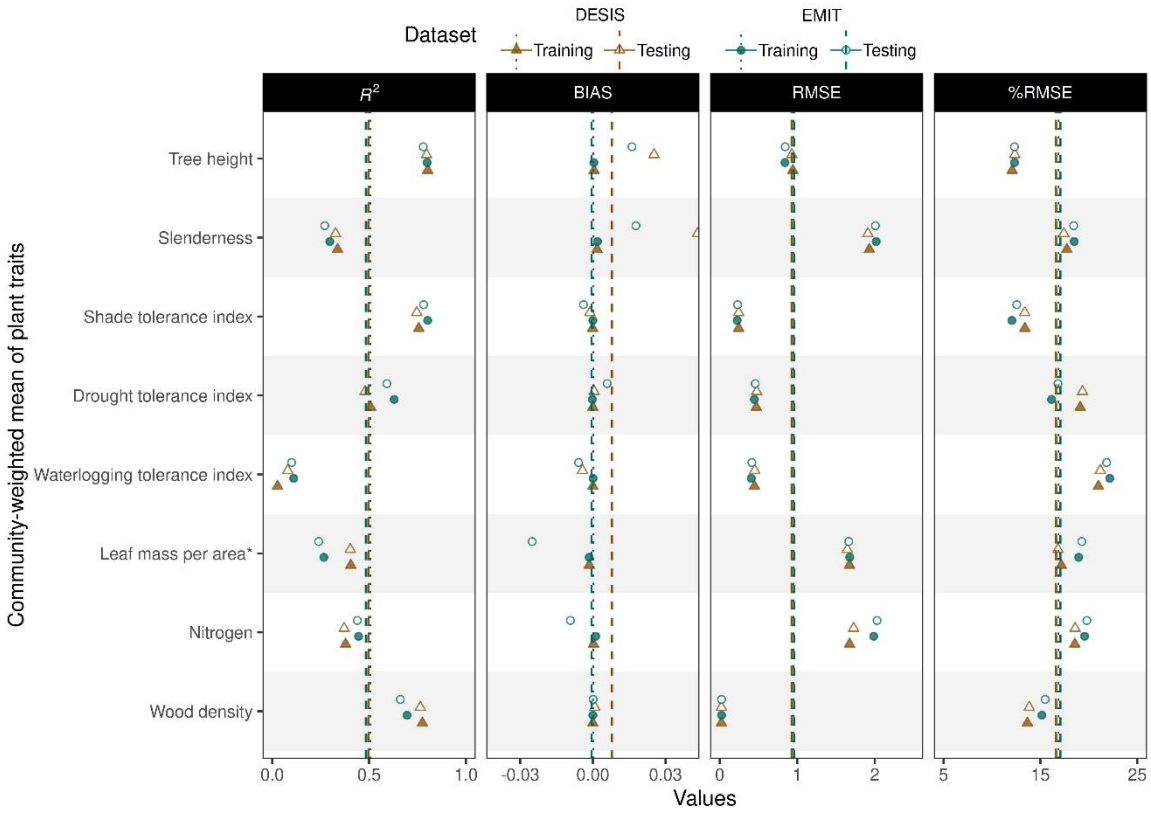


Fig. S32. Performance to predict the community-weighted mean of plant traits within inventories of the Forest Inventory and Analysis program using climatic and topographic information as well as predicted MDS axes of dimensions of beta diversity based on spaceborne observations of DESIS or EMIT. Each point represents the mean estimate of 100 models.

References

Debastiani, V.J., Bastazini, V.A.G., Pillar, V.D., 2021. Using phylogenetic information to impute missing functional trait values in ecological databases. Ecol. Inform. 63, 101315. https://doi.org/10.1016/j.ecoinf.2021.101315

EROS, 2017. Global Topographic 30 Arc-Second Hydrologic Digital Elevation Model 1 km. https://doi.org/10.5066/F77P8WN0

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315. https://doi.org/10.1002/joc.5086

Hargrove, W.W., Hoffman, F.M., 1999. Using multivariate clustering to characterize ecoregion borders. Comput. Sci. Eng. 1, 18–25. https://doi.org/10.1109/5992.774837

Maitner, B.S., Boyle, B., Casler, N., Condit, R., Donoghue, J., Durán, S.M., Guaderrama, D., Hinchliff, C.E., Jørgensen, P.M., Kraft, N.J.B., McGill, B., Merow, C., Morueta-Holme, N., Peet, R.K., Sandel, B., Schildhauer, M., Smith, S.A., Svenning, J., Thiers, B., Violle, C., Wiser, S., Enquist, B.J., 2018. The BIEN R package: A tool to access the Botanical Information and Ecology Network (BIEN) database. Methods Ecol. Evol. 9, 373–379. https://doi.org/10.1111/2041-210X.12861

Niinemets, Ü., Valladares, F., 2006. Tolerance to shade, drought, and waterlogging of temperate Northern Hemisphere trees and shrubs. Ecol. Monogr. 76, 521–547. https://doi.org/10.1890/0012-9615(2006)076%5B0521:TTSDAW%5D2.0.CO;2

Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20, 289–290. https://doi.org/10.1093/bioinformatics/btg412

Santos, T., 2018. PVR: Phylogenetic Eigenvectors Regression and Phylogentic Signal-Representation Curve.

signal developers, 2023. signal: Signal processing.

Smith, S.A., Brown, J.W., 2018. Constructing a broadly inclusive seed plant phylogeny. Am. J. Bot. 105, 302–314. https://doi.org/10.1002/ajb2.1019

Stekhoven, D.J., Bühlmann, P., 2012. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics 28, 112–118. https://doi.org/10.1093/bioinformatics/btr597