Disentangling the roles of inter and intraspecific variation on leaf trait distributions across the eastern United States

Sergio Marconi[1], Ben Weinstein[1], Jeremy Lichstein[1], Stephanie Bohlman[1], Aditya Singh[1], and Ethan White[1]

[1]University of Florida

January 28, 2022

Abstract

Functional traits are influenced by phylogenetic constraints and environmental conditions, but previous large-scale studies modeled traits either as species weighted averages or directly from the environment, precluding analyses of the relative contributions of inter- and intraspecific variation across regions. We developed a joint model integrating phylogenetic and environmental information to understand and predict the distribution of eight leaf traits across the eastern US. This model explained 68% of trait variation, outperforming both species-only and environment-only models, with variance attributable to species alone (23%), the environment alone (13%), and their overlapping effects (25%). The importance of the two drivers varied by trait. Predictions for the eastern US produced accurate estimates of intraspecific variation and deviated from both species-only and environment-only models. Predictions revealed that intraspecific variation holds information across scales, affects relationships in the leaf economic spectrum and is key for interpreting trait distributions and ecosystem processes within and across ecoregions.

Introduction

Global change is expected to cause extensive changes in terrestrial ecosystems, driving unprecedented redistribution of plant species and their associated traits (Pecl et al., 2017, Diaz and Cabido, 2001). Plant functional traits are involved in key ecosystem processes from local community assembly (McGill et al., 2006, Sterck et al., 2011) to global biogeochemical cycles, and these processes are interconnected across scales (Reichstein et al., 2014, Peaucelle et al., 2019). Relationships between traits, such as the leaf economic spectrum (LES)(Wright et al., 2005) , reveal information about biological constraints in leaf mass allocation that impact plant ecophysiology and have the potential to improve ecosystem models (Fisher et al, 2015). Biotic interactions, micro-climate, and soil conditions can affect species co-occurrence and influence local trait distributions (Bruelheide et al., 2018, Simpson et al., 2016), and variation in climate within species ranges can affect realized niches and drive trait responses (Chave, 2013). Given their central role across levels of organization, understanding how traits vary within and among species across scales and environments is essential for conserving present and future ecosystem function (Violle et al., 2014).

Plant traits vary geographically through a combination of interspecific shifts in species abundances and intraspecific trait variation (Leps et al., 2011, Laughlin et al., 2012, Valladares et al., 2014, Münzbergová et al., 2017). Understanding how traits respond to the environment across wide geographic areas requires approaches integrating both sources of trait variation. This is challenging because individual level trait data are geographically and taxonomically limited, making it hard for traditional methods to identify the relative importance of inter- and intraspecific variation at large scales (Henn et al., 2018). In addition, effectively predicting broad-scale geographic patterns requires making predictions for species not included in trait datasets but widely distributed across the continent. Most current approaches to understanding and predicting trait variation use community weighted mean approaches (CWM), which circumvent these limitations by either focusing directly on trait–environment relationships (based on direct relationships between environment and

1

community trait averages while ignoring species) or by estimating traits from species averages (using the environment only for predicting species assembly but ignoring the effect of intraspecific variation) (Miller & Ives, 2019).

CWMs relying on direct trait-environment relationships offer the advantage of predicting large scale community trait distributions without requiring field surveys or estimates of species assembly. However, by ignoring species identity (e.g., Ordonez et al., 2009) this approach ignores known phylogenetic signals in trait variation driven by biological, physical, and historical constraints (Wright et al., 2004, Anderegg et al., 2018), and assumes that the environment implicitly captures relevant changes in species distribution and abundance. Since the information about species identity is missing from predictions, these models cannot be used to explicitly identify the relative contributions of inter- and intraspecific variation (but see Moles et al. 2014) and could potentially yield worse predictions for assemblage level trait values (when abundance data is available) unless the environmental model fully captures relevant shifts in species abundance.

Due to these limitations, it has been suggested that predicting traits directly from species' average values (Swenson 2010, Clark 2016, Wieczynski et al., 2019, Swenson 2017, Stahl et al., 2014) offers better estimates of large-scale trait distributions. These approaches assume that environmental drivers affect trait distributions indirectly by shaping community structure and species abundance, implying that species distributions are the best predictor of traits and associated ecosystem function. These models can be used to make predictions for traits over large areas by training on forest inventories, leveraging species distribution models to forecast future shifts in species and consequently trait distributions (Swenson & Weiser, 2010, Clark 2016). Yet, this approach ignores intraspecific variation, which can be larger than interspecific variation for broadly distributed species (Niinemets, 2015, Messier et al., 2017), and overlooks that species averages within regional communities may diverge from their global averages (Hulshof & Swenson, 2010).

Both these approaches contribute to our understanding of trait variation and allow for predicting community level trait distributions without requiring extensive field surveys of traits. However, neither approach is designed to predict trait variation at the individual level, nor allow for estimating intra-species variation. Also, they often fail to account for the effect of phylogenetic signal on traits of closely related species (but see Swenson et al., 2017), which can be potentially important for generating more robust trait predictions for species that are sparsely sampled (or not sampled at all) across vast geographic areas (Blomberg et al. 2003, Swenson 2013, Swenson et al., 2014). These limitations prevent assessment of the relative importance of inter- and intraspecific variation on trait distributions across a continuum of geographical scales, reducing our ability to generalize and understand mechanisms driving trait distributions.

To address these limitations, we developed a model that combines species, their phylogenetic relationships (from the Tree of Life; Hinchliff et al., 2015) and environmental drivers (climate, elevation, slope, terrain aspect) with large scale leaf trait data from the National Ecological Observatory Network (NEON) (National Ecological Observatory Network, 2020). While this method can only predict trait distributions for locations with species abundance data, it makes it possible to estimate the relative contribution of intra and interspecies variation on trait distributions across wide geographic areas. This allows us to address whether changes in environmental conditions have a direct effect on the LES, within and across species. We jointly modeled eight leaf traits: nitrogen (N%), carbon (C%), chlorophyllA (ChlA%), chlorophyllB (ChlB%), carotenoids (Crt%), leaf mass per area (LMA, g m^{-2}), lignin (%) and cellulose (%). We compared this combined model to models based on only environmental drivers or only species and phylogeny information. We integrated the combined model with US Forest Inventory and Analysis (FIA; USDA Forest Service, 2001, Smith et al., 2002) and Daymet data (Thornton et al., 2018) to make trait predictions for ~1.2 million trees across the eastern US. We compared these predictions to the other two approaches to assess the influence of model differences on large scale prediction, analyzed the relative contribution of environmental factors and phylogeny to leaf trait variation across ecoregions, and demonstrate the potential use of this data for understanding the processes structuring ecological systems at scale.

Materials and Methods

Data

We used data from the National Ecological Observatory Network (NEON, National Ecological Observatory Network. 2020), the Botanical Information and Ecology Network (BIEN, Maitner et al., 2020, Enquist et al., 2009) and TRY (Kattge et al., 2020) to link information on leaf traits, species identity, and approximate locations for individual trees. We used Foliar Physical and Chemical Properties (DP1.10026.001) and Vegetation Structure data (DP1.10098.001) from NEON to build joint trait distribution models with environmental drivers (climate and topography) alone, phylogenetic drivers (species identity and phylogeny) alone, and both (combined model). Linking the two different NEON datasets produced individual tree data with stem geolocation and measures of eight leaf traits (LMA, chlorophyll A and B, carotenoids, lignin, cellulose, C, N) for 542 trees in 21 sites across the US (Figure S.1). Since foliar trait concentrations vary significantly with phenology and canopy position (Niinemets et al., 2015), foliar samples were collected at the "peak of greenness" and from the sunlit portion of the canopy. We tested the generalizability of our approach outside of NEON by evaluating predictions from independent (out of sample) data available from the BIEN and TRY datasets (Appendix S1). These two datasets provide measures for C, N and LMA for a total of 223 individual trees. We used data from the Open Tree of Life (Redeling,s 2017) to measure phylogenetic distance between species.

Data for environmental drivers included average monthly climate data from 1995 to 2015 (Appendix S1) extracted from Daymet (Thornton et al., 2018) and topographic variables (elevation, slope and aspect) reported in the NEON and FIA datasets. For three common eastern US tree species (*Acer rubrum* , *Fagus grandifolia* , and *Abies balsamea*), we used all publicly available leaf N% data from the TRY database to quantify intraspecific variation in leaf N% across each species' geographic range in the US. We selected these three species because: (1) *Abies balsamea* is the needleleaf species with the most leaf N% data in TRY for the US; (2) *Fagus grandifolia* is the broadleaf species with the most leaf N% data in TRY for the USA; and (3) *Acer rubrum* occurs throughout much of the eastern US in a wide variety of habitats (e.g., from xeric to mesic; Burns and Honkala 1990) and has abundant leaf N% data in TRY. We combined our trait modeling approach with forest survey data from the Forest Inventory and Analysis (FIA) database (*https://www.fia.fs.fed.us/*) to estimate traits for all individual trees surveyed in the FIA across the eastern USA from 2016 to 2019. We used Lv.3 ecoregions and Lv. 2 ecoprovinces as defined by the Environmental Protection Agency (McMahon 2001, Omernik et al., 2014) to analyze trait distributions at different scales.

Overview of Models

We modeled the joint multivariate distribution of the eight leaf traits (the response variables) using three different approaches: (1) Environment-only model using climate and topography as fixed effects; (2) Phylogeny-only model using species as random effects, with covariances among the random effects structured by the phylogenetic tree for all woody species detected in the FIA database for the eastern US; (3) Combined model including both environmental and phylogenetic effects. Here, we briefly summarize the modeling framework; full details are in Appendix S2. All approaches used the same joint-multilevel Bayesian framework and were evaluated using 5-fold cross-validation. The joint structure of the model allows for modeling traits simultaneously, considering the correlation structure across traits embedded in the LES and potentially leveraging the conditional distribution on known trait values of one or more individuals at a site (Wilkinson et al., 2020). Environmental effects were fitted using generalized additive models (GAMs) to account for non-linear relationships. We used thin plate regression splines to estimate the smooth terms, using the brms R package (Bürkner, 2017). Phylogenetic relationships across species were modeled by including species as a random effect and accounting for their phylogenetic relationships by estimating their correlation structure from cross-species cophenetic distances (Paradis et al., 2019). The distance matrix was used to estimate the correlation structure across taxa, allowing parameter estimates for rare or unsampled taxa to borrow strength from widely sampled species (de Villemeruil & Nakagawa, 2014). We used multivariate normal families and weakly informative priors in all cases (Appendix S2).

To reduce collinearity and the number of climate predictors, we calculated a PCA for each climate variable (net radiation, precipitation, vapor pressure, maximum and minimum temperature) using monthly avera-

ges from 1985 to 2015. We used the first component of each PCA to represent each climate variable in the environment-only and combined models. To quantify uncertainty in model accuracy, we used the 95% prediction interval of the Bayesian R^2 (Gelman et al., 2018). To reduce computation costs of ~15x without affecting accuracy (Table S.1), we used predictions from 1fold cross-validation for making predictions across FIA and test model generalizability. See additional methods details in Appendix S1-S3. Code for reproducing analyses is available on Zenodo (*https://zenodo.org/badge/latestdoi/353383665*).

Estimation of inter- intraspecific variation across scales

We used variance partitioning among two factors (Ribas et al. 2006) to estimate the relative contribution of inter- vs. intraspecific trait variation at the continental scale by comparing each model (phylogenetic and environmental) to the combined model following Munoz & Real (2006). The idea of using variance partitioning to understand intra- vs inter-specific trait variation was originally developed by de Bello et al. (2011; see also Leps et al. 2006), but this work focused on estimating variance from field data in the absence of a model. We use the traditional approach from statistics by quantifying the proportion of variance attributable to purely phylogenetic drivers ($\sigma^2_{Phylo} = R^2_{comined} - R^2_{Env}$), purely environmental drivers ($\sigma^2_{Env} = R^2_{comined} - R^2_{Phylo}$), and the proportion that is jointly shared between them ($\sigma^2_{joint} = R^2_{comined} - \sigma^2_{Phylo} - \sigma^2_{Env}$).

To quantify the relative effects of environment, species and phylogeny on inter- and intraspecific trait variation at the regional scale (eastern US), we used the combined model to estimate the leaf traits for all individual trees in the FIA data.

Results and Discussion

Model evaluation

We built the phylogeny-only, environment-only, and combined (phylogeny and environment) models using leaf trait data from NEON and evaluated their explanatory power using the Bayesian R^2 of the predicted values for 88 out-of-sample test trees. For all 8 leaf traits, the combined model explained the largest amount of variance in the held-out test data (average R^2 across the 8 traits = 0.64), substantially outperforming both the environment-only (average R^2 = 0.35) and phylogeny-only (average R^2 = 0.52) models (Figure 1, Figure S.2). A hierarchical clustering of model residuals supported two major trait classes (Figure S.3): traits mainly involved in photosynthesis (Croft et al., 2017) and traits involved in leaf structure. The combined model had the highest performance for predicting LMA (R^2 = 0.81) and the lowest performance for predicting ChlA (R^2 = 0.51). Uncertainty in predictions was accurately estimated across all traits and for all models (Figure S.4), with the combined model showing mean 95% coverage values ranging from 94.3% to 98.8%. The importance of different environmental drivers varied among traits, supporting an important role of climate in driving leaf economics in local communities (Ordoñez et al. 2009). Precipitation and temperature were mainly important for traits involved in photosynthesis (N%, ChlA, ChlB and Carotenoids), generally having a positive effect on their concentration (except for ChlB, Figure S.5). Net radiation showed a negative effect on N%, while vapor pressure generally had a negative effect on pigments but a positive effect on traits associated with leaf toughness and durability (cellulose, lignin and C%). Elevation was the most important topographic predictor and the strongest environmental driver of LMA (Figure S.5), consistent with previous studies (Reich & Oleksyn, 2004, Hedin 2004, Poorter et al., 2009, Kitajima et al., 2016). The joint model structure also captured the strong correlation among LMA and N% characteristic of the leaf economics spectrum (Reich et al, 1997, Wright et al., 2004).

Role of species and environment in predicting traits at the continental scale

We used variance partitioning on the Bayesian R^2 from out-of-sample test data to explore the relative contributions of inter- and intraspecific trait variation at continental scale (Supplement 4). On average, interspecific variation (pure phylogenetic and species effect) accounted for 25% of the total explained variation across the 8 traits, intraspecific variation (pure environment effect) accounted for 13%, and joint phylogenetic-environment effects accounted for 23%. The relative importance of inter- and intraspecific effects varied widely among traits. Species and phylogeny explained most of the variation in structural traits (e.g. LMA, C% and

4

lignin%). For these traits, often used in large scale distribution studies, species distributions may contain more information about traits than direct predictions from the environment, as previously suggested by CWM models (Clark 2016, Yang & Swenson, 2018). These results are aligned with previous studies suggesting that the relative extent of intra-species variation among communities is negatively related to spatial extent (Siefert, et al. 2015) and that at continental and global scales, patterns of leaf traits are mainly driven by leaf economic strategies at the species level (Wright et al., 2005). In contrast, the intraspecific component of the model accounted for as much or more of the variance than the interspecific one for pigments.

This difference between structural and photosynthesis traits may be explained by being driven by different kinds of tradeoffs. Structural traits are more affected by leaf lifespan and toughness, which varies widely across species (Wright et al. 2004, Kitajima et al. 2012, Osnas et al. 2018, Lichstein et al. 2021), whereas photosynthetic traits are less variable among species due to the fundamental need for all species to maximize carbon gain (subject to ecological conditions and tradeoffs; Wright et al. 2004; Maire et al, 2015). Our analysis only quantifies intraspecific variation from the upper (sun-lit) canopy. Although this is a common practice (Pérez-Harguindeguy et al. 2013), it likely results in underrepresenting intraspecific variation by ignoring leaf variability across the light gradient, a major source of intraspecific variation (Osnas et al., 2018). Furthermore, 62% of species in our analysis were only sampled within a single NEON site, thus representing only a small fraction of species' true environmental ranges. Thus, our analysis likely underestimates the true level of intraspecific trait variation.

Model transferability

To determine the appropriateness of using the model to make predictions outside of the scope of the NE-ON dataset, we tested the performance of the combined model on novel locations and novel species using independent data from the Botanical Information and Ecology Network (BIEN) and a subset of TRY leaf trait datasets. These datasets include trait data on LMA, N% and C% for 62 species, including 27 species unavailable in training data (Figure S.1, supplement S1). The combined model showed good transferability to other data sources (mean $R^2 = 0.54$, 95% coverage = 91%, Figure S.6). The inclusion of full phylogenetic relationships in addition to environmental predictors yielded successful model transfer to species not sampled at NEON sites (mean $R^2 = 0.4$, 95% coverage = 94%, Figure 2). This was possible because phylogenetic relationships allow parameter estimates for unsampled species to borrow strength from closely related species included in the data (Evans et al., 2016). The model is therefore suitable for large scale application.

Predicting large scale trait variation

To understand large scale variation in traits using our combined model and to compare it to the phylogeny-only and environment-only approaches, we integrated each of the three models with tree species abundance and topographic data from ~30,000 Forest Inventory and Analysis (FIA) plots (~1.2 million trees) and climate data from DayMet (Figure S.7, S.8, S.9) to predict leaf traits across the eastern US. Predictions from the phylogeny-only model produced patterns similar to Swenson & Weiser (2010) (Figure S.10), suggesting that our approach to representing species-only methods (and any resulting deviations from them) is consistent with patterns previously reported in literature.

Predictions from the combined model show broad-scale patterns associated with shifts in forest communities and large-scale climatic and topographic patterns across latitudinal and altitudinal gradients (Figure 3, S.11). In some cases, trait distributions shifted abruptly between neighboring ecoregions due to a combination of shifts in local environmental conditions and in community assembly (independent of the environment). Changes in species composition may explain trait patterns between the Mississippi Plains ecoregion and Southern Plains ecoregions (Figure S.12). The Mississippi Plains ecoregion is characterized by heavy disturbance from agricultural activities and forests are often limited to riparian ecosystems favoring bottomwood broadleaf species (e.g., *Celtis laevigata* , *Fraxinus pennsylvanica* ,*Salix nigra*) in contrast to needleleaf species (e.g.,*Juniperus virginiana* , *Pinus taeda* , and *Pinus echinata*) more common in the neighboring ecoregions (Coastal plains mixed forests). Being that broadleaf species are generally characterized by higher N% and lower LMA, this change in community assembly translates into predicted regional trait patterns. Other

5

ecoregions show little change in species composition compared to neighboring ecoregions instead, suggesting that shifts in predicted patterns may be attributed to how environmental gradients affect traits directly. This seems to be the case of the mixed forests in the Appalachian region, where opposing patterns are exhibited at higher altitudes in the Blue Ridge ecoregion (lower N% and pigment concentrations; higher LMA and C%) compared to piedmont and valleys in the neighboring ecoregions (higher N% and pigments, lower LMA and C%) (Figure S.13).

Predictions from the combined model differed from the phylogeny- and environment-only models, suggesting that phylogeny and environmental drivers contain different information at large scales. Divergence from the combined model varied across ecoregions (Figure 4). These differences were complex, with regions exhibiting shifts of different magnitudes and directions from either phylogeny- or environment-only approaches (Figure 4, Figure S.14, Supplement 5). 80% of ecoregion-trait-model combinations exhibited significant differences in predicted traits between the combined and phylogeny- or environment-only models ($p < 0.0001$ in paired t-tests). Accordingly, predictions from the phylogeny-only and environment-only model differed from each other ($p < 0.0001$ in paired t-tests) for 93% of ecoregion-trait combinations, which highlights the distinct effects of the environment and phylogeny on trait distributions and demonstrates the importance of a combined modeling approach for prediction and inference.

Patterns of divergence from the combined model indicate how phylogenetic and environmental effects vary biogeographically across the eastern US. Regions with no significant divergence may signal conditions where the environment affects traits by filtering for species better adapted to local conditions, and traits are well estimated by both species' averages and directly from the environment (as in the case of the Mississippi Alluvial Plains region). For most ecoregions, divergence of the phylogeny- and environment-only models move in opposite directions (i.e., positive divergence for one opposed to negative divergence for the other). Significant divergence for the phylogeny-only model indicates that continental-scale species averages fail to correctly represent trait values at finer scales, because local environmental conditions and/or competition shift the trait values away from the species mean.

Areas where the combined model predicts higher N% than the phylogeny-only model (blue shading in Fig. 4a) suggest the presence of environmental effects that increase N% above species means. Conversely, orange-shaded areas in Fig. 4a indicate environmental effects that lower N% below species means. These patterns may reflect regulatory mechanisms controlled by the environment, where leaves adjust allocation to proteins, pigments, or structural compounds to balance photosynthetic capacity, toughness, and chemical defense within ranges constrained by species life history (Weih & Karlsson, 2001, Tjoelker et al., 2001, Crous et al., 2019, Albert et al., 2010). In contrast, divergence between the combined and environment-only models (Fig. 4b) could be explained by (i) environmental effects that have a phylogenetic signal not captured by the environmental variables included in our analysis; and/or (ii) stochastic factors (e.g., disturbance history and dispersal limitation) that have resulted in species distribution patterns that are decoupled from current environmental conditions (Burns & Strauss, 2012, McIntyre et al., 1999). Testing mechanistic hypotheses for the above divergence patterns is beyond the scope of our study. Nevertheless, merely identifying these biogeographic patterns is a novel step towards better understanding trait distributions and is only possible using modeling approaches that combine phylogenetic and environmental information.

Large scale intraspecific trait variation

In addition to mapping phylogenetic and environmental trait signals, another essential reason for a combined approach is that it allows for predictions of intraspecific variation (Figure 5). Models based on CWMs, by definition, do not account for intraspecific variation. Although environment-only models implicitly include intraspecific variation, it is not decoupled from trait variation due to shifts in species composition (Hulshof & Swenson, 2010). This interspecific variation must be separated from intraspecific variation to quantify the latter. Our combined modeling approach makes this separation possible for hundreds of species-trait combinations. Quantifying intraspecific variation in a comprehensive manner (i.e., for all species and across a large geographic region) is a necessary step towards testing hypotheses about the role of individual trait variation in species coexistence and interspecific competition (Hart et al., 2016). We explore the potential

for our approach to enable research understanding the role of intraspecific variation in community dynamics by (i) comparing our estimates of intraspecific variation to field data for three widespread species, (ii) describing how the predicted intraspecific variation varies among eastern US species, and (iii) describing how trait variation may affect trends in one of the trait-tradeoffs described by the LES.

The combined model produced realistic ranges of intraspecific variation when compared with available independent data for widely distributed species. Specifically, the intraspecific variation of N% predicted from the combined model showed ranges similar to those observed from the field by independent datasets for three widespread and abundantly sampled species in the NEON, FIA and TRY datasets (*Abies balsamea, Acer rubrum* , and *Fagus grandifolia*). For these species (Figure 5b-d), predictions from the combined model showed N% covering a large range of values (1.12 N% on average), comparable to the average difference between evergreen needleleaf and deciduous broadleaf species in the eastern US (1.13 N% based on NEON field data). Such a wide range of intra-species variation further supports the idea that adaptation to the environment accounts for a large proportion of community-level variation and may be driving community-level shifts across environmental gradients (Albert et al. 2010, Violle et al. 2012, Siefert et al. 2015, Fajardo & Siefert, 2018). The contribution of the environment to intra-species variation may also explain altitudinal (e.g. *Acer rubrum)* and latitudinal (e.g. *Abies balsamea*) gradients observed for these three species.

Predicted patterns of intraspecific variation varied widely across species. For N%, the ratio of predicted intraspecific to total observed interspecific variation ranged from less than 10% for species with limited geographic ranges (e.g., *Populus heterophylla* ,*Sabal palmetto* , and *Gleditsia aquatica*) to over 60% for broadleaf species with broad geographic ranges (e.g., *Cercis canadensis, Betula lenta* , and *Carpinus caroliniana*). Intraspecific responses of N% to temperature also showed a variety of patterns among species (Figure 5a, Figure S.15), including 1) bell-shaped patterns (n=85) like those observed in compilations of field data (Reich & Oleksyn, 2004, Laughlin et al., 2012); 2) negative relationships (n = 7); and 3) cases with no significant relationship (n=108). We observed these different patterns despite the model not including species-by-environment interactions. This is possible due to differences in how the environmental drivers are jointly related in different subregions of the posterior predictive distribution that represent different geographic regions and different environments by trait combinations. Therefore, while the underlying relationship between traits and temperature in the model contains a single dominant mode with an uptick at very high temperatures, analyzing patterns of trait distributions from predictions across large geographic areas shows that focusing only on the subregions of the posterior that are relevant to the species of interest can yield different species level relationships.

Variation in environmental drivers may also impact trends in trait trade-offs, including those described in the LES. Since these relationships affect nutrient and carbon use efficiency in plants (Reich, 2014), understanding how they change geographically under different environmental conditions is fundamental for improving how carbon dynamics are simulated by earth system models (Weng et al., 2017). Our results suggest that within the Eastern US, the relationship between N% and LMA is weaker than the global inter-species average, but stronger than local within-species relationships (Figure 6). This relationship, first observed from field data from the western US (Anderegg et al., 2018), generalizes to the Eastern US and holds when extrapolating trait values continuously across species ranges. The scale-dependent change in the strength of the relationship may be driven by different nitrogen to mass allocation strategies between maximizing short term productivity (higher foliar photosynthetic mass) or long-term defense (higher foliar structural mass) (Osnas et al., 2018). According to this conceptual model, we expect stronger negative N%-LMA relationship for datasets with high variation in LMA, due to variation in leaf structural components. This is the case of the global LES, calculated on species sampled from all biomes and many ecosystems (slope of the logarithmic N%-LMA linear relationship ~-0.7, Wright et al., 2004). At local scales, where macro-environmental drivers have a much smaller effect on LMA variation, adaptations to local conditions may drive higher variation in photosynthetic mass leading to weaker N% to LMA relationships (slope of the logarithmic N%-LMA linear relationship ~-0.2; Osnas et al. 2018). At the intermediate scale of the Eastern US, both strategies may contribute significantly to the relationship, resulting in N%-LMA trends falling halfway in between the two extremes (slope of the logarithmic N%-LMA linear relationship ~-0.43). This slope was consistent across broadleaves, needleleaves,

deciduous and evergreen species, providing support that this trait-tradeoff can be represented the same across plant functional types used in earth system models.

Implications for estimating processes from Earth System Models

These results have important implications for the use of earth system models (ESMs) for estimating eco-systems processes across scales. Most ESMs treat traits as constants, disregarding variation among species within plant functional types and intraspecific variation driven by the environment (Ghimire et al., 2016, Lawrence et al., 2019, but see attempts at integrating more flexible strategies like Fisher et al., 2015). This was justified by previous evidence suggesting that only interspecific variation mattered at continental scales (Messier et al. 2010). However, our results demonstrate that even at near-continental scales intraspecific variation holds meaningful information that needs to be addressed (Figures 1, 5). The need to incorporate intraspecific variation is magnified when models are applied at smaller scales where the environment affects traits-tradeoffs to move away from the global LES (Fisher et al., 2018). Because different regions exhibit different directional shifts from the species-only models (Figure 4) we need to understand environmentally driven intraspecific traits variation to accurately model regional ecosystem processes.

Conclusions

Both phylogenetic and environmental effects are fundamental to understanding the drivers and distribution of plant traits. Combining both in a single model is challenging due to data limitations but is possible by leveraging large scale datasets. This approach allows for improved traits predictions compared to models that rely on either species average or the environment in isolation and allows for robust predictions for species and regions not included in the training data. Across Eastern US both interspecies trait variation (driven by shifts in species' abundance) and intraspecific variation are key for predicting joint trait distributions, with effects on the LES. The influence of these components varies by species, trait, ecoregion and scale.

Our approach overcomes previous data limitations by integrating multiple sources of biological and environmental information to create a single integrated model. As new traits, phylogenetic, and species inventory data is released globally, the combined approach can be extended to new regions and unlock the potential to study patterns of intraspecific variation for hundreds of traits-species-environment combinations. For example, this is already possible by leveraging national forest inventories in some European countries, Canada, New Zealand (Schelhaas et al, 2006, Rati et al., 2018, Gills et al., 2005, Paul et al., 2021) along with ever-growing plant traits datasets stored in TRY (*https://www.try-db.org/*). Expanding this work outside the US could contribute to further understanding the mechanisms driving trait distributions across scales and the link between traits, species distribution, forest assembly and function.

Acknowledgements

Description of author's responsibilities: Sergio Marconi, Ben Weinstein and Ethan White designed the experiment and developed the methods; Sergio Marconi performed the analysis; Ethan White, Jeremy Lichstein, Aditya Singh and Stephanie Bohlman supervised the work, helped with experimental design and advice on data analysis and technical aspects. All authors contributed to editing the manuscript.

References

1. Albert, C.H., Thuiller, W., Yoccoz, N.G., Soudant, A., Boucher, F., Saccone, P. and Lavorel, S., 2010. Intraspecific functional variability: extent, structure and sources of variation. Journal of Ecology, 98(3), pp.604-613.
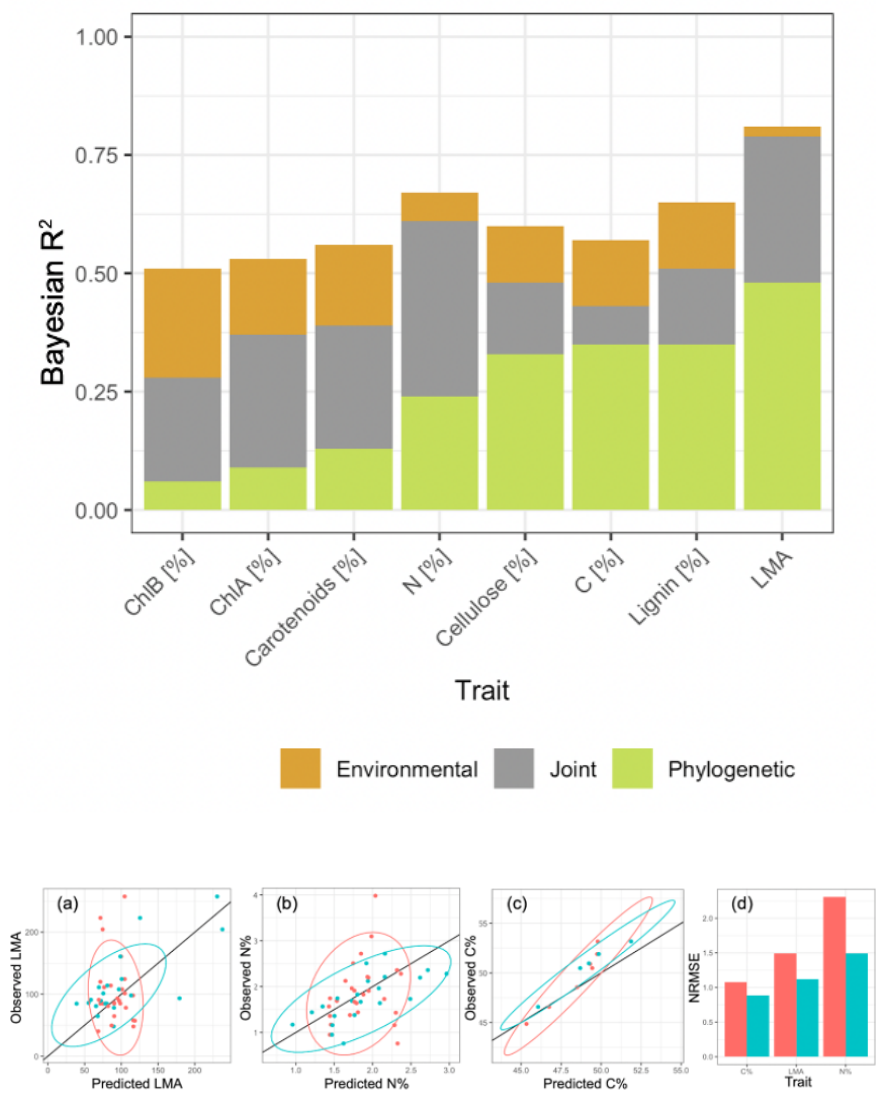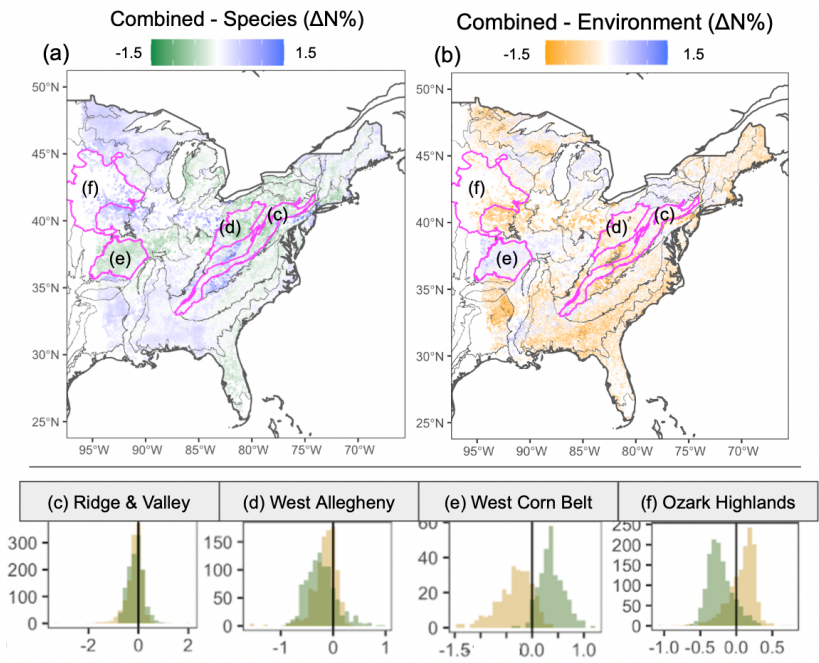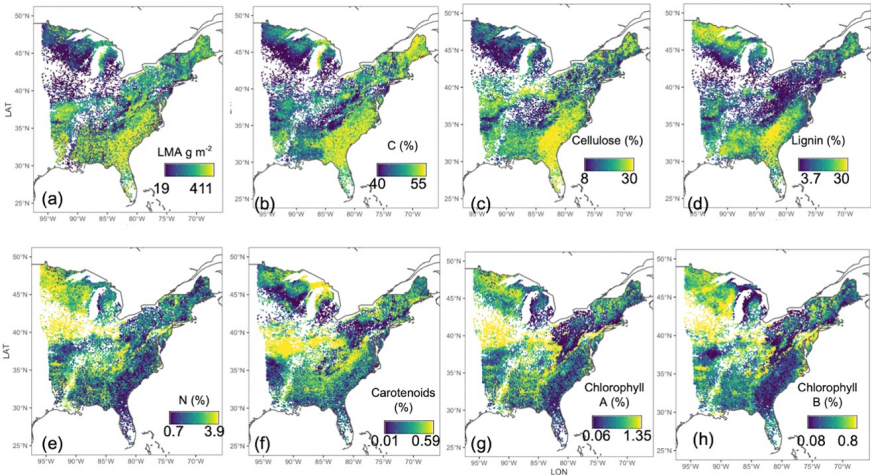
2. Anderegg, L. D., Berner, L. T., Badgley, G., Sethi, M. L., Law, B. E., & HilleRisLambers, J. (2018). Within-species patterns challenge our understanding of the leaf economics spectrum. Ecology letters, 21(5), 734-744.

3. Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jimenez-Alfaro, B., Hennekens, S.M., Botta-Dukat, Z., Chytry, M., Field, R., Jansen, F. and Kattge, J., 2018. Global trait–environment relationships of plant communities. Nature ecology & evolution, 2(12), p.1906.

4. Burkner P.C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80(1), 1-28. doi:10.18637/jss.v080.i01

5. Burns, R.M. & Honkala, B.H. (1990). Silvics of North America: 1. Conifers; 2. Hardwoods. Agriculture Handbook 654. U.S. Department of Agriculture, Forest Service, Washington, DC.

6. Burns, J. H., & Strauss, S. Y. (2012). Effects of competition on phylogenetic signal and phenotypic plasticity in plant functional traits. Ecology, 93(sp8), S126-S137.

7. Butler, E.E., Datta, A., Flores-Moreno, H., Chen, M., Wythers, K.R., Fazayeli, F., Banerjee, A., Atkin, O.K., Kattge, J., Amiaud, B. and Blonder, B., 2017. Mapping local and global variability in plant trait distributions. Proceedings of the National Academy of Sciences, 114(51), pp.E10937-E10946.

8. Chave, J., 2013. The problem of pattern and scale in ecology: what have we learned in 20 years?. Ecology letters, 16, pp.4-16.

9. Clark, J. S. (2016). Why species tell more about traits than traits about species: predictive analysis. Ecology, 97(8), 1979-1993.

10. Croft, H., Chen, J.M., Luo, X., Bartlett, P., Chen, B. and Staebler, R.M. (2017), Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. Glob Change Biol, 23: 3513-3524. doi:10.1111/gcb.13599

11. Crous, K. Y. (2019). Plant responses to climate warming: physiological adjustments and implications for plant functioning in a future, warmer world. American journal of botany, 106(8), 1049.

12. de Bello, F., et al. 2011. Quantifying the relevance of intraspecific trait variability for functional diversity. - Methods in Ecology and Evolution 2: 163-174.

13. De Villemeruil P. & Nakagawa, S. (2014) General quantitative genetic methods for comparative biology. In: Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice(ed. Garamszegi L.) Springer, New York. pp. 287-303.

14. Diaz, S., and M. Cabido. 2001. Vive la difference: plant functional diversity matters to ecosystem processes. Trends in Ecology & Evolution 16:646–655.

15. Enquist, B.J., Condit, R., Peet, R.K., Schildhauer, M. and Thiers, B., 2009. The Botanical Information and Ecology Network (BIEN): Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. Salt Lake City, UT: iPlant Collaborative.

16. Evans, M.E., Merow, C., Record, S., McMahon, S.M. and Enquist, B.J., 2016. Towards process-based range modeling of many species. Trends in Ecology & Evolution, 31(11), pp.860-871.

17. Fisher, Rosie A., et al. "Vegetation demographics in Earth System Models: A review of progress and priorities." *Global change biology* 24.1 (2018): 35-54.

18. Fisher, R. A., Muszala, S., Verteinstein, M., Lawrence, P., Xu, C., McDowell, N. G., ... & Bonan, G. (2015). Taking off the training wheels: the properties of a dynamic vegetation model without climate envelopes, CLM4. 5 (ED). Geoscientific Model Development, 8(11), 3593-3619.

19. Gelman, Ben Goodrich, Jonah Gabry & Aki Vehtari. (2018). R-squared for Bayesian regression models, The American Statistician. https://doi.org/10.1080/00031305.2018.1549100.

20. Ghimire, B., Riley, W. J., Koven, C. D., Mu, M., & Randerson, J. T. (2016). Representing leaf and root physiological traits in CLM improves global carbon and nitrogen cycling predictions. Journal of Advances in Modeling Earth Systems, 8(2), 598-613.

21. Gillis, Mark D., A. Y. Omule, and T. Brierley. "Monitoring Canada's forests: the national forest inventory." The Forestry Chronicle 81.2 (2005): 214-221.

22. Hart, S. P., et al. 2016. How variation between individuals affects species coexistence. - Ecology Letters 19: 825-838.

23. Hedin, L.O., 2004. Global organization of terrestrial plant–nutrient interactions. Proceedings of the National Academy of Sciences, 101(30), pp.10849-10850.

24. Henn, J.J., Buzzard, V., Enquist, B.J., Halbritter, A.H., Klanderud, K., Maitner, B.S., Michaletz, S.T., Potsch, C., Seltzer, L., Telford, R.J. and Yang, Y., 2018. Intraspecific trait variation and phenotypic plasticity mediate alpine plant species response to climate change. Frontiers in Plant Science, 9, p.1548.

25. Hinchliff, Cody E., et al. "Synthesis of phylogeny and taxonomy into a comprehensive tree of life." Proceedings of the National Academy of Sciences 112.41 (2015): 12764-12769. https://doi.org/10.1073/pnas.1423041112

26. Kattge, J., Bonisch, G., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Tautenhahn, S., Werner, G.D., Aakala, T., Abedi, M. and Acosta, A.T., 2020. TRY plant trait database–enhanced coverage and open access. Global change biology.

27. Kitajima K, Wright SJ, Westbrook JW (2016) Leaf cellulose density as the key determinant of inter- and intra-specific variation in leaf fracture toughness in a species-rich tropical forest. Interface Focus 6:20150100.

28. Lamanna, C., Blonder, B., Violle, C., Kraft, N.J., Sandel, B., Šímová, I., Donoghue, J.C., Svenning, J.C., McGill, B.J., Boyle, B. and Buzzard, V., 2014. Functional trait space and the latitudinal diversity gradient. Proceedings of the National Academy of Sciences, 111(38), pp.13745-13750.

29. Laughlin, D. C., et al. 2012. A predictive model of community assembly that incorporates intraspecific trait variation. - Ecology Letters 15: 1291-1299.

30. Lepš, J., et al. 2011. Community trait response to environment: disentangling species turnover vs intraspecific trait variability effects. - Ecography 34: 856-863.

31. Lichstein, J. W., Peterson, B. T., Langebrake, J., & McKinley, S. A. (2021). Leaf economics of early- and late-successional plants.

32. Maire, V., Wright, I. J., Prentice, I. C., Batjes, N. H., Bhaskar, R., Van Bodegom, P. M., . . . & Santiago, L. S. (2015). Global effects of soil and climate on leaf photosynthetic traits and rates. Global Ecology and Biogeography, 24(6), 706-717.

33. Maitner B. (2020). BIEN: Tools for Accessing the Botanical Information and Ecology Network Database. R package version 1.2.4. https://CRAN.R-project.org/package=BIEN

34. McGill, B.J., Enquist, B.J., Weiher, E. and Westoby, M., 2006. Rebuilding community ecology from functional traits. Trends in ecology & evolution, 21(4), pp.178-185.

35. McIntyre, S., Díaz, S., Lavorel, S., & Cramer, W. (1999). Plant functional types and disturbance dynamics–Introduction. Journal of Vegetation Science, 10(5), 603-608.

36. McMahon, G., Gregonis, S. M., Waltman, S. W., Omernik, J. M., Thorson, T. D., Freeouf, J. A., . . . & Keys, J. E. (2001). Developing a spatial framework of common ecological regions for the conterminous United States. Environmental Management, 28(3), 293-316.

37. Miller, J. E., Damschen, E. I., & Ives, A. R. (2019). Functional traits and community composition: A comparison among community-weighted means, weighted correlations, and multilevel models. Methods in Ecology and Evolution, 10(3), 415-425.

38. Moles, A.T., Perkins, S.E., Laffan, S.W., Flores-Moreno, H., Awasthy, M., Tindall, M.L., Sack, L., Pitman, A., Kattge, J., Aarssen, L.W. and Anand, M., 2014. Which is a better predictor of plant traits: temperature or precipitation?. Journal of Vegetation Science, 25(5), pp.1167-1180.

39. Munzbergova, Z., Hadincova, V., Skalova, H. and Vandvik, V., 2017. Genetic differentiation and plasticity interact along temperature and precipitation gradients to determine plant performance under climate change. Journal of Ecology, 105(5), pp.1358-1373.

40. Munoz A.-R. & Real R. (2006) Assessing the potential range expansion of the exotic monk parakeet in Spain. Diversity and Distributions 12: 656-665.

41. NEON (National Ecological Observatory Network), 2020. Plant foliar traits (DP1.10026.001). Provisional data downloaded from http://data.neonscience.org on 2 Jan 2020. Battelle, Boulder, CO, USA

42. Niinemets, U., Keenan, T.F. and Hallik, L., 2015. A worldwide analysis of within-canopy variations in leaf structural, chemical and physiological traits across plant functional types. New Phytologist,

205(3), pp.973-993.

43. Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the conterminous United States: evolution of a hierarchical spatial framework. Environmental management, 54(6), 1249-1266.

44. Ordonez, Jenny C., et al. "A global study of relationships between leaf traits, climate and soil measures of nutrient fertility." Global Ecology and Biogeography 18.2 (2009): 137-149.

45. Osnas, J.L., Katabuchi, M., Kitajima, K., Wright, S.J., Reich, P.B., Van Bael, S.A., Kraft, N.J., Samaniego, M.J., Pacala, S.W. and Lichstein, J.W., 2018. Divergent drivers of leaf trait variation within species, among species, and among functional groups. Proceedings of the National Academy of Sciences, 115(21), pp.5480-5485.

46. Paradis, E., Blomberg, S., Bolker, B., Brown, J., Claude, J., Cuong, H.S. and Desper, R., 2019. Package 'ape'. Analyses of phylogenetics and evolution, version, 2(4).

47. Peaucelle, M., Bacour, C., Ciais, P., Vuichard, N., Kuppel, S., Penuelas, J., Belelli Marchesini, L., Blanken, P.D., Buchmann, N., Chen, J. and Delpierre, N., Covariations between plant functional traits emerge from constraining parameterization of a terrestrial biosphere model. Global Ecology and Biogeography.

48. Pecl, G.T., Araujo, M.B., Bell, J.D., Blanchard, J., Bonebrake, T.C., Chen, I.C., Clark, T.D., Colwell, R.K., Danielsen, F., Evengard, B. and Falconi, L., 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science, 355(6332), p.eaai9214.

49. Perez-Harguindeguy, N., et al. 2013. New handbook for standardised measurement of plant functional traits worldwide. - Australian Journal of Botany 61: 167-234.

50. Poorter, H., Niinemets, U., Poorter, L., Wright, I.J. and Villar, R., 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. New phytologist, 182(3), pp.565-588.

51. Raty, Minna, Juha Heikkinen, and Annika Kangas. "Assessment of sampling strategies utilizing auxiliary information in large-scale forest inventory." Canadian Journal of Forest Research 48.7 (2018): 749-757.

52. Redelings BD, Holder MT. "A supertree pipeline for summarizing phylogenetic and taxonomic information for millions of species." PeerJ. 2017;5:e3058. https://doi.org/10.7717/peerj.3058

53. Reich, Peter B. "The world-wide 'fast–slow' plant economics spectrum: a traits manifesto." Journal of Ecology 102.2 (2014): 275-301.

54. Reich, P.B., Oleksyn., J., 2004. Global patterns of plant leaf N and P in relation to temperature and latitude. Proceedings of the National Academy of Sciences, 101 (30) 11001-11006; DOI: 10.1073/pnas.0403588101

55. Reich, P.B., Walters, M.B. & Ellsworth, D.S. (1997). From tropics to tundra: Global convergence in plant functioning. Proc. Natl. Acad. Sci. U.S.A., 94, 13730–13734.

56. Reichstein, M., Bahn, M., Mahecha, M.D., Kattge, J. and Baldocchi, D.D., 2014. Linking plant and ecosystem functional biogeography. Proceedings of the National Academy of Sciences, 111(38), pp.13697-13702.

57. Ribas A., Barbosa A.M., Casanova J.C., Real R., Feliu C. & Vargas J.M. (2006) Geographical patterns of the species richness of helminth parasites of moles (Talpa spp.) in Spain: separating the effect of sampling effort from those of other conditioning factors. Vie et Milieu 56: 1-8.

58. Schelhaas, M.J., Varis, S., Schuck, A. and Nabuurs, G.J., 2006, EFISCEN Inventory Database, European Forest Institute, Joensuu, Finland, http://www.efi.int/portal/virtual_library/databases/efiscen/

59. Siefert, A., et al. 2015. A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. - Ecology Letters 18: 1406-1419.

60. Simpson, A.H., Richardson, S.J. and Laughlin, D.C., 2016. Soil–climate interactions explain variation in foliar, stem, root and reproductive traits across temperate forests. Global Ecology and Biogeography, 25(8), pp.964-978.

61. Smith, W.B., 2002. Forest inventory and analysis: a national inventory and monitoring program. Environmental pollution, 116, pp.S233-S242.

62. Stahl, U., Reu, B. and Wirth, C., 2014. Predicting species' range limits from functional traits for the tree flora of North America. Proceedings of the National Academy of Sciences, 111(38), pp.13739-

13744.

63. Sterck, F., Markesteijn, L., Schieving, F. and Poorter, L., 2011. Functional traits determine trade-offs and niches in a tropical forest community. Proceedings of the National Academy of Sciences, 108(51), pp.20627-20632.

64. Swenson, N.G., Weiser, M.D., Mao, L., Araujo, M.B., Diniz-Filho, J.A.F., Kollmann, J., Nogues-Bravo, D., Normand, S., Rodriguez, M.A., Garcia-Valdes, R. and Valladares, F., 2017. Phylogeny and the prediction of tree functional diversity across novel continental settings. Global Ecology and Biogeography, 26(5), pp.553-562.

65. Swenson, N.G., 2014. Phylogenetic imputation of plant functional trait databases. Ecography, 37(2), pp.105-110.

66. Swenson, N.G., 2013. The assembly of tropical tree communities–the advances and shortcomings of phylogenetic and functional trait analyses. Ecography, 36(3), pp.264-276.

67. Swenson, N. G., & Weiser, M. D. (2010). Plant geography upon the basis of functional traits: an example from eastern North American trees. Ecology, 91(8), 2234-2241.

68. Tjoelker, M. G., Oleksyn, J., & Reich, P. B. (2001). Modelling respiration of vegetation: evidence for a general temperature-dependent Q10. Global change biology, 7(2), 223-230.

69. Thornton, M.M., P.E. Thornton, Y. Wei, B.W. Mayer, R.B. Cook, and R.S. Vose. 2018. Daymet: Annual Climate Summaries on a 1-km Grid for North America, Version 3. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1343

70. USDA Forest Service, 2001. Forest Inventory and Analysis National Core Field Guide, Volume I: Field Data Collection Procedures For Phase 2 Plots, Version 1.5. US Department of Agriculture, Forest Service, Washington, DC.

71. Valladares, F., Matesanz, S., Guilhaumon, F., Araujo, M.B., Balaguer, L., Benito-Garzon, M., Cornwell, W., Gianoli, E., van Kleunen, M., Naya, D.E. and Nicotra, A.B., 2014. The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. Ecology letters, 17(11), pp.1351-1364.

72. Violle, C., Reich, P.B., Pacala, S.W., Enquist, B.J. and Kattge, J., 2014. The emergence and promise of functional biogeography. Proceedings of the National Academy of Sciences, 111(38), pp.13690-13696.

73. Weih, M., & Karlsson, P. S. (2001). Growth response of Mountain birch to air and soil temperature: is increasing leaf-nitrogen content an acclimation to lower air temperature?. New Phytologist, 150(1), 147-155.

74. Weng, Ensheng, Caroline E. Farrior, Ray Dybzinski, and Stephen W. Pacala. "Predicting vegetation type through physiological and environmental interactions with leaf traits: evergreen and deciduous forests in an earth system modeling framework." Global change biology 23, no. 6 (2017): 2482-2498.

75. Wieczynski, D.J., Boyle, B., Buzzard, V., Duran, S.M., Henderson, A.N., Hulshof, C.M., Kerkhoff, A.J., McCarthy, M.C., Michaletz, S.T., Swenson, N.G. and Asner, G.P., 2019. Climate shapes and shifts functional biodiversity in forests worldwide. Proceedings of the National Academy of Sciences, 116(2), pp.587-592.

76. Wilkinson, D. P., Golding, N., Guillera-Arroita, G., Tingley, R., & McCarthy, M. A. (2021). Defining and evaluating predictions of joint species distribution models. Methods in Ecology and Evolution, 12(3), 394-404.

77. Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H., Diemer, M. and Flexas, J., 2004. The worldwide leaf economics spectrum. Nature, 428(6985), p.821.

78. Yang, J., Cao, M. and Swenson, N.G., 2018. Why functional traits do not predict tree demographic rates. Trends in ecology & evolution, 33(5), pp.326-336

79. Zhang, Y.J., Sack, L., Cao, K.F., Wei, X.M. and Li, N., 2017. Speed versus endurance tradeoff in plants: Leaves with higher photosynthetic rates show stronger seasonal declines. Scientific reports, 7, p.42085.

Combined - Species (ΔN%)

Combined - Environment (ΔN%)

Divergence from Combined model (ΔN%)