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

Sergio Marconi[1,2], Benjamin G. Weinstein[2], Jeremy W. Lichstein[3], Stephanie A. Bohlman[1], Aditya Singh[4], Ethan P. White[2]

[1]School of Forest, Fisheries, & Geomatics Sciences, University of Florida, Gainesville, Florida, USA

[2]Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, Florida, USA

[3]Department of Biology, University of Florida, Gainesville, Florida, USA

[4]Department of Agricultural & Biological Engineering, University of Florida, Gainesville, FL 32611, USA

Running Title: Inter-intraspecies traits variation across east USA

Statement of authorship: SM, BW and EW designed the experiment and developed the methods; SM performed the analysis; EW, JL, AS and SB supervised the work, helped with experimental design and supervised data analysis and technical aspects. All authors contributed to editing the manuscript.

Data accessibility: All data and data products are archived and accessible on Zenodo. Code to replicate the analyses and retrieve data is stored on GitHub.

Keywords: leaf traits, intra-species variability, joint distribution, cross-scale, NEON, FIA.

Words in the abstract: 150

Number of Figures: 6

Number of references: 79

Competing Interest Statement: No conflict of interests.

Authors email:

Sergio Marconi: s.marconi@ufl.edu

Benjamin G. Weinstein: ben.weinstein@weecology.org

Jeremy Lichstein: jlichstein@ufl.edu

Stephanie Bohlman: sbohlman@ufl.edu

Aditya Singh: aditya01@ufl.edu

Ethan White: ethanwhite@ufl.edu

Corresponding Author: Sergio Marconi

Email: s.marconi@ufl.edu

Address: School of Forest, Fisheries, and Geomatics Sciences, University of Florida, 136 Newins-Ziegler Hall, Gainesville, FL 32611

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 USA. 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 combined effects (25%). The importance of the two drivers varied by trait. Predictions for the eastern USA 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 (Bonan et al., 2002; Verheijen et al., 2013), and these processes are interconnected across

scales (Reichstein et al., 2014, Peaucelle et al., 2019). 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). Likewise, relationships between traits such as those captured by the leaf economic spectrum (LES) (Wright et al., 2004), reveal information about biological constraints in leaf mass allocation that impact plant ecophysiology and are key for improving ecosystem models (Fisher et al, 2015). 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).

Accurately representing leaf functional diversity in terrestrial ecosystem models (e.g., biogeochemical and/or land-surface models) requires capturing variation within and among species (Bonan et al., 2002, Thornton and Zimmermann 2007, Osnas et al., 2018). The LES provides a simple, one-dimensional framework for understanding leaf functional variation from local to global scales (Wright et al., 2004) and for representing plant functional diversity in terrestrial ecosystem models (Sakschewski et al., 2015), including the land components of Earth System Models (Bonan et al., 2002, Fisher et al., 2015). On the other hand, divergent patterns in trait variation observed within species, among species, and among functional or taxonomic groups (Osnas et al., 2018) and at different taxonomic scales (Anderegg et al., 2018) suggest that a one-dimensional framework may be too simplistic. There is growing interest in enhanced representations of trait diversity in terrestrial ecosystem models (Scheiter et al., 2013, Wullschleger et al., 2014, Fisher et al., 2015, Sakschewski et al., 2015, Dantas de Paula et al.,

2021), but these model developments are currently data-limited due to limited information on trait distributions at broad geographic scales (see details below).

Plant traits vary geographically through a combination of shifts in species abundances and intraspecific trait variation (Leps et al., 2011, Laughlin et al., 2012, Valladares et al., 2014, Münzbergová et al., 2017). Thus, understanding how traits respond to the environment across wide geographic areas requires approaches that integrate both inter- and intraspecific 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, due to limited data availability, studying traits at broad geographic scales may require predicting trait values for species that have not been sampled throughout their geographic range, or even species that have not been sampled at all. Given these challenges, most current approaches to studying broad-scale trait variation focus on predicting community weighted mean (CWM) trait values, which circumvents data limitations by either focusing directly on trait–environment relationships (based on direct relationships between environment and CWMs without explicitly considering species distributions or abundances) or by estimating CWMs directly from species mean trait values (without accounting for intraspecific variation, and considering environmental variables only for the purpose of predicting species distributions) (Miller & Ives, 2019).

Models based on direct trait-environment relationships offer the advantage of predicting trait distributions without requiring data on species abundances. A corollary of this approach is that

it ignores well-established species and phylogenetic signals in trait variation driven by biological, physical, and historical constraints (Wright et al., 2004, Ordonez et al., 2009, Anderegg et al., 2018), and implicitly assumes that the environment captures relevant changes in species distribution and abundance. However, since species-environment relationships are typically noisy (Clark 2016), failing to explicitly account for species distribution and abundance inevitably introduces prediction errors. Furthermore, models that do not explicitly incorporate species effects cannot be used to study inter- and intraspecific contributions to trait variation, which has important implications for functional ecology, ecophysiology, and ecosystem modeling (Anderegg et al., 2018, Osnas et al., 2018, Tautenhahn et al., 2020).

An alternative to pure environmental models is to predict traits directly from species-level means, which may provide better estimates of large-scale trait distributions (Swenson 2010, Clark 2016, Wieczynski et al., 2019, Swenson 2017, Stahl et al., 2014). Species-based approaches assume that environmental drivers affect trait distributions only indirectly by shaping community assembly, but that species distributions themselves are the best predictor of traits and associated ecosystem function. Species-based models can be used to make predictions for traits over large areas by leveraging national-scale forest inventories or other vegetation plot networks, and can be combined with species distribution models to forecast future shifts in trait distributions (Swenson & Weiser, 2010, Clark 2016). However, species-level approaches ignore intraspecific variation, which can be larger than interspecific variation for broadly distributed species (Niinemets, 2015, Messier et al., 2017).

Both pure environment and pure species approaches have contributed to our understanding of trait variation and allow for predicting community level trait distributions without requiring

extensive field surveys of traits. However, as noted above, neither approach is designed to study intraspecific trait variation: pure environment approaches cannot separate inter- and intraspecific variation, and pure species approaches completely ignore intraspecific variation. Also, these approaches often fail to account phylogenetic signals among closely related species (but see Swenson et al., 2017), which is potentially important for generating more robust trait predictions for species that are sparsely sampled, or not sampled at all (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 and inter-trait relationships, such as those embodied in the LES.

To address these limitations, we developed a model that combines species identity (including phylogenetic relationships from the Tree of Life; Hinchliff et al., 2015) and environmental drivers (climate and topography) with leaf trait data from the National Ecological Observatory Network (NEON) (National Ecological Observatory Network, 2020). Our approach makes it possible to estimate the relative contributions of inter- and intraspecific variation to trait distributions at any location or scale where species abundance data are available. This allows us to address whether changes in environmental conditions have a direct effect on trait distributions and inter-trait relationships, 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 our combined species-environment model to models based on only environmental drivers or only on species and phylogenetic information. We integrated the combined model with Forest

Inventory and Analysis (FIA) data (Bechtold, W. A., and P. L. Patterson. 2005) and Daymet data (Thornton et al., 2018) to make leaf trait predictions for ~1.2 million trees across the eastern USA and validated the predictions against independent field datasets.

We compared predictions from our combined model to predictions from environment-only and species-only approaches to assess the influence of inter- vs. intraspecific variation on trait distributions across scales, species, and functional types. Specifically, we present four applications, in which we: (1) quantify the relative contributions of environmental factors vs. species and phylogenetic information to leaf trait variation at continental scale; (2) assess drivers of ecoregional differences in trait distributions; (3) analyze large-scale intraspecific trait variation driven by environmental drivers; and (4) explore how intraspecific trait variation impacts the relationship between N% and LMA (one of the LES relationships) and the implications for representing leaf trait variation in ecosystem models.

Materials and Methods

Data

We used data from 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 477 trees in 20 sites across the USA (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) eastern USA data available from the BIEN and TRY datasets (Appendix S1). These two datasets provide measures for C, N and LMA for a total of 353 individual trees. We used data from the Open Tree of Life (Redelings 2017) to measure phylogenetic distances among species, which we used to inform the covariance structure of our combined and species-only models.

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 USA 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 USA. We selected these three species because: (1) *Abies balsamea* is the needleleaf species with the most leaf N% data in TRY in the USA; (2) *Fagus grandifolia* is the broadleaf species with the most leaf N% data in TRY in the USA; and (3) *Acer rubrum* occurs throughout much of the eastern USA 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 data from the USA Forest Inventory and Analysis

(FIA) database (https://www.fia.fs.usda.gov/) 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 define analysis regions 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) species-only model using species as random effects, with covariances among the random effects structured by the phylogeny for all woody species recorded in the FIA database for the eastern USA; and (3) 'combined model' including both environmental and species effects (throughout this paper, 'species effects' in our models refers to phylogenetically structured species random 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 all traits simultaneously (Wilkinson et al., 2020), considering the correlation structure across traits (i.e., inter-trait relationships, such as those typical of the LES). 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 the lead traits of 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 performed a PCA for each climate variable (net radiation, precipitation, vapor pressure, maximum and minimum temperature) using monthly averages 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). See additional methods details in Appendix S1-S3. Code for reproducing our analyses is available on Zenodo (https://zenodo.org/badge/latestdoi/353383665).

Partitioning trait variation

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 single-driver model (environment-only and species-only) to the combined model following Muñoz & Real (2006). Previous studies partitioned inter- vs. intraspecific trait variation using measured trait values (Leps et al., 2006; de Bello et al., 2011). To adapt this framework to the modeling context, we quantified the proportion of variance explained (Bayesian R^2) for each trait by each of our three models (environment, species, and combined models), and we then quantified the proportion of variance attributable to species effects alone ($\sigma^2_{species} = R^2_{combined} - R^2_{env}$), environmental effects alone ($\sigma^2_{env} = R^2_{combined} - R^2_{species}$), and joint species-environment

effects ($\sigma^2_{joint} = R^2_{combined} - \sigma^2_{species} - \sigma^2_{env}$). Note that the pure environment effect (σ^2_{env}) is a component of intraspecific variation, because σ^2_{env} is, by definition, independent of species-level (interspecific) effects ($\sigma^2_{species}$). However, not all intraspecific variation can be attributed to measured environmental variables; thus, intraspecific variation contributes not only to σ^2_{env} , but also to unexplained variation. In summary, trait variation can be conceptualized as including pure species effects ($\sigma^2_{species}$), pure environmental effects (σ^2_{env}, which is equivalent to intraspecific variation explained by the measured environmental variables), joint species-environment effects (σ^2_{joint}), and unexplained variation (which includes measurement error as well as intraspecific variation associated with errors or omissions in the environmental data).

Analysis methods for quantifying inter- and intraspecific variation at scale

To study inter- and intraspecific patterns in trait variation, including inter-trait relationships, we used the combined model to estimate leaf traits for all individual trees in FIA plots in the eastern USA (all USA states bordering or to the east of the Mississippi River). For each individual tree (~1.2 million trees in ~30,000 inventory plots), we estimated its posterior median trait values using each of our three models. For the environment-only and combined models, we used the climate principal components (see above) at the publicly reported location of each FIA plot, and we used the topographic variables (elevation, slope, and aspect) reported by FIA for each plot. Predictions from our species-only model produced broad-scale geographic patterns similar to those from previous species-only models (Swenson & Weiser, 2010; Clark, 2016) (Figure S.2). We compared these predictions from our species-only model (which ignores intraspecific variation) to predictions from our combined model (which includes both inter- and

intraspecific variation) to explore the consequences of intraspecific variation for broad-scale trait distributions.

To quantify relationships between N% and LMA, we used ordinary least squares (OLS) regression slopes based on log-transformed predictions of N% and LMA from our combined model for individual trees in FIA plots.. To quantify differences in N% vs. LMA relationships within and among species and plant functional types, we compared the OLS slopes from regressions based on individuals vs. species-level means, for all eastern USA species and separately for broadleaf and needleleaf plant functional types. We used OLS regression instead of standard major axis (SMA) regression because (1) this facilitated comparisons with the OLS slopes reported by Osnas et al., (2018), including inter- and intraspecific trait relationships for different species groups; (2) SMA slopes are unstable when OLS slopes are shallow, as observed for a large fraction species; and (3) we expect only small errors in the x-variable (LMA) in these regressions (see additional details on the N% vs. LMA regressions in Supplement 6).

Results and Discussion

Model evaluation: explained variation at continental scale

We fit the species-only, environment-only, and combined (species 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 out-of-sample test data (average R^2 across the 8 traits = 0.64), substantially outperforming both the environment-only (average R^2 = 0.35) and species-only (average R^2 = 0.52) models (Figure 1, Figure S.3). The combined model had the

highest performance for predicting LMA (R^2 = 0.81) and the lowest performance for predicting ChlBA (R^2 = 0.51). Uncertainty in predictions was accurately estimated across all traits and for all models (Figure S.4), with the combined model's mean 95% coverage values ranging from 94.3% to 98.8%. Residuals from the combined model – which represent variation that is unexplained by environmental factors, species identity, or phylogeny – revealed patterns similar to those previously described across species, suggesting the presence of inherent trait trade-offs and correlations that occur not only across species, but also within species in a given environment. Specifically, combined-model residuals revealed a strong negative correlation between N% and LMA, as in the LES (Reich et al, 1997, Wright et al., 2004), and the presence of two main trait clusters (Figure S.5): traits mainly involved in photosynthesis (N%, Carotenoids, ChlA and ChlB; Croft et al., 2017) and traits related to leaf structure (LMA, C%, Cellulose%, Lignin%; Lima et al., 2018).

Model evaluation: geographic transferability

To determine the appropriateness of using the model to make predictions outside of the scope of the NEON 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 TRY. These datasets include trait data on LMA, N% and C% for 62 eastern USA tree 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). Accounting for phylogenetic relationships, in addition to

environmental predictors, yielded substantial model improvements when predicting traits for species not sampled at NEON sites (Figure 2). This improvement occurs because accounting for phylogenetic structure when modeling species random effects (see Methods) allows parameter estimates for unsampled species to borrow strength from closely related species for which observations are available (Evans et al., 2016). The above transferability analyses suggest that the combined model is suitable for broad-scale applications.

Model evaluation: intraspecific variation across wide geographic ranges

The combined model produced realistic ranges of intraspecific variation when compared with available independent data for widely distributed species (Figure 3). Specifically, intraspecific ranges of N% predicted from the combined model were similar to those observed in field samples 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 3b-d), combined-model predictions of N% spanned an average range of 1.12 N%, which is as wide as the average difference between evergreen needleleaf and deciduous broadleaf species in the eastern USA (1.13 N% based on NEON field data). This large magnitude of intraspecific variation supports the idea that adaptation to the environment accounts for a large proportion of community-level variation and plays an important role in community-mean trait shifts across environmental gradients (Albert et al., 2010, Violle et al., 2012, Siefert et al., 2015, Fajardo & Siefert, 2018).

Application 1: Role of species and environment in predicting traits at the continental scale

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.7). 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, consistent with previous studies (Reich & Oleksyn, 2004, Hedin 2004, Poorter et al., 2009, Kitajima et al., 2016).

On average, interspecific variation (phylogenetically structured species-level effects, controlling for environmental variation; see $\sigma^2_{species}$ in Methods) accounted for 25% of the total explained variation across the 8 traits, intraspecific variation (pure environmental effects, controlling for species effects; see σ^2_{env} in Methods) accounted for 13%, and joint species-environment effects accounted for 23% (Figure 1). However, variance partitioning differed for photosynthetic traits (pigments and N%) vs. structural traits (cellulose%, C%, lignin%, and LMA). For photosynthetic traits, the largest variance component was the joint species-environment component (Figure 1); i.e., for these traits, much of the explained variation was associated with coordinated shifts in the environment and species composition. In contrast, for structural traits, pure species effects (i.e., interspecific variation that is independent of environmental variation) accounted for the majority of trait variation (Figure 1). It is important to note, however, that data limitations likely lead us to underestimate the importance of pure environmental (intraspecific variation) effects. Firstly, our analysis is restricted to the upper (sun-lit) canopy. Although this is

a common practice (Pérez-Harguindeguy et al., 2013), it ignores trait variation across the canopy light gradient, a major source of intraspecific variation (Lusk et al., 2008; Osnas et al., 2018). Secondly, 62% of species in our analysis were only sampled within a single NEON site. Thirdly, as in any observational study, we have only limited environmental data, which inevitably leads to an underestimation of environmentally structured intraspecific variation. Thus, our analysis likely underestimates the true level of intraspecific trait variation both within and across sites.

Despite the likely underestimation of intraspecific variation in our analysis, pure environmental effects (a component of intraspecific variation; see Methods) accounted for a substantial fraction of explained variation for all traits except for N% and LMA (Figure 1). Thus, although species-only models accounted for the majority of explained variation (i.e., the sum of the joint effects and species effects in Figure 1) – consistent with previous studies highlighting the important role of species distributions in predicting traits at broad geographic scales (Wright et al., 2005; Siefert, et al., 2015; Clark 2016, Yang & Swenson, 2018) – predictions for most traits can be substantially improved by a combined species-environment modeling approach. This conclusion – i.e., that accounting for intraspecific variation can enhance trait predictions − would likely be strengthened by improved data on traits and environmental variables.

Application 2: Drivers of ecoregional differences in traits distribution

Predictions from the combined model show broad-scale patterns associated with shifts in forest communities (Figure S.8) and large-scale climatic and topographic patterns across latitudinal

and altitudinal gradients (Figure 4). 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). For example, changes in species composition may explain trait differences between the Mississippi Alluvial Plains and neighboring ecoregions (Figure S.9). The Mississippi Alluvial Plains ecoregion is characterized by agricultural land use, and forests are often limited to riparian ecosystems favoring bottomland 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. 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 ecoregion boundaries show little change in species composition, in which case geographic variation in traits may be attributed to environmental effects on intraspecific variation. For example, in the Appalachian Mountains, opposing patterns are exhibited at the 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.10).

Predictions from the combined model differed from the species-only and environment-only models, with the direction and magnitude of divergences varying across ecoregions (Figure 5, Figure S.11). Species-only and environment-only model predictions differed significantly from the combined model ($p < 0.0001$ in paired t-tests) in 80% cases (different combinations of traits, ecoregions, and models). Similarly, predictions from the species-only and environment-only models differed from each other ($p < 0.0001$ in paired t-tests) for 93% of ecoregion-trait

combinations, which demonstrates the distinct effects of species and environmental drivers on trait distributions and underscores the importance of a combined modeling approach for prediction and inference.

Patterns of divergence from the combined model indicate how species and environmental effects vary geographically across the eastern USA (Figure 4). In regions with no significant divergence (e.g., the Ridge and Valley region), species-only and environment-only models yield similar results to each other and to the combined model. This convergence suggests a limited role for intraspecific variation combined with strong environmental filtering of species assembly. However, for most ecoregions, species-only and environment-only models diverge from the combined model in opposite directions (i.e., positive divergence for one and negative divergence for the other) (Figure 5d, 5e, 5f, S.11). Significant divergence for the species-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 species-only model (blue shading in Figure 5a) suggest the presence of environmental effects that increase N% above species means. Conversely, orange-shaded areas in Figure 5a indicate environmental effects that lower N% below species means. These patterns may be explained by regulatory mechanisms adjusting allocation to proteins, pigments, or structural compounds to balance photosynthetic capacity, toughness, and chemical defense across environmental gradients (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 (Figure 5b) 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, identifying these biogeographic patterns is a step towards better understanding trait distributions and is only possible using modeling approaches that combine phylogenetic and environmental information.

Application 3: Large scale intraspecific trait variation driven by environmental drivers

In addition to mapping species and environmental trait signals, another essential reason for a combined approach is that it allows for predictions of intraspecific variation for individual species (Figure 3). Models based on community-weighted means, by definition, do not account for intraspecific variation. In contrast, environment-only models implicitly capture at least some component of intraspecific variation (see Methods) but cannot separate this intraspecific variation from trait variation due to shifts in species composition (Hulshof & Swenson, 2010). 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 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). Here, we explore the potential for using a combined species-environment modeling

approach to studying intraspecific variation by describing how predicted levels of intraspecific variation vary among eastern USA 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 3a, Figure S.12), 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). Note that our modeling approach does not explicitly include species-by-environment interactions. Thus, differences among species in their magnitudes of intraspecific variation and in their trait-environment relationships emerge from how species distributions in environmental space combine with non-linear trait-environment relationships.

Application 4: Inter- and intraspecific trait relationships: implications for the leaf economics spectrum and terrestrial ecosystem models

When expressed per-unit leaf mass, leaf traits related to photosynthesis, metabolism, and nutrient concentrations typically have strong negative covariance with LMA across species (Wright et al., 2004, 2005, Lloyd et al., 2013, Osnas et al., 2013) but are often roughly invariant within species across canopy light gradients (Ellsworth and Reich 1993, Evans and Poorter 2001,

Thornton and Zimmermann 2007, Niinemets et al., 2015). Constant mass-normalized traits across canopy light gradients − equivalent to a roughly linear increase in the area-normalized value of these same traits as LMA increases from shade to sun (Osnas et al., 2018) − is accounted for in some terrestrial ecosystem models (e.g., Sellers et al., 1992, Thornton and Zimmermann 2007, Bonan et al., 2012). Intraspecific variation in leaf traits across broad environmental gradients can be substantial as well, with important consequences for ecosystem models (Reich et al., 2014). Our broad-scale systematic mapping of N% and LMA, for individual trees (summarized at the FIA plot level in Figure 4, Figure S.13) enables further understanding of intraspecific trait relationships across geographic space.

For the eastern USA, our results show that slopes of N% vs LMA within and among species fall between the shallow slopes found within species across light gradients (typically between 0 and −0.3; Supplement 6), and the steeper slopes recorded among global species (−0.6; Supplement 6). These results were consistent for both deciduous broadleaf (DB) and evergreen needleleaf (EN) plant functional types (PFTs; Table S.2). Specifically, in our analysis, intraspecific slopes had median values around −0.37 (Figure 6a-b), slopes across all individuals within PFTs were -0.36 for DB and -0.48 for EN (Figure 6c), and slopes across species means were -0.4 for DB and -0.45 for EN (Figure 6d). These intermediate values (i.e., steeper than those observed across intraspecific light gradients, but shallower than observed across global species) may be explained by partially compensating factors. LMA variation associated with intraspecific LL variation across space may largely involve variation in low-nutrient, structural components of LMA (as occurs across species; Kitajima et al., 2012), which would lead to steep N% vs. LMA slopes (Osnas et al., 2018). LMA variation associated with certain climatic responses (e.g.,

sensitivity of leaf water potential to tissue water content; Niinemets 2001) may also largely involve low-nutrient structural mass, which would again lead to steep N% vs. LMA slopes. In contrast, water loss can be reduced at dry sites (which tend to have high LMA) if N% and thus photosynthetic efficiency are also high (Wright et al., 2001), which would lead to positive N% vs. LMA slopes across moisture gradients. This combination of mechanisms, and perhaps similar mechanisms operating along other geographic gradients (e.g., season length and temperature extremes), may together result in intraspecific leaf trait relationships across space that are intermediate to those driven by photosynthetic responses to light vs. those driven by interspecific variation in LL.

Our results suggest that to accurately represent leaf functional variation, terrestrial ecosystem models may need to consider three different types of variation: intraspecific variation across light gradients, intraspecific variation across geographic environmental gradients, and interspecific relationships (which may differ among PFTs). Given the impracticality of parameterizing species-level global ecosystem models, we expect these models to continue to explore approaches to representing trait variation within broadly defined PFTs without explicitly modeling individual species (e.g., Fisher et al., 2015, Sakschewski et al., 2015). Accurately representing the different intra- and interspecific patterns described above does not necessarily require an explicit species-level approach. A greater understanding of the anatomical and physiological mechanisms underlying trait variation (e.g., Shipley et al., 2006, Sack et al., 2006, 2013, Sack and Scoffoni 2013, Niinemets et al., 2015, John et al., 2017) could lead to a generalized, species-independent model that captures divergent patterns within and among species and PFTs.

Finally, our results have implications for interpreting interspecific leaf trait relationships. PFT-level slopes across individuals (Figure 6c) differed from slopes across species means (Figure 6d). These differences in slope are greatly reduced when all species are rarified to equal abundance (Figure S.14), suggesting that the slope differences are largely due to differences in species abundance: more common species exert greater influence in the individual-level analysis (Figure 6c), whereas all species contribute equally to the species-level analysis (Figure 6d). In at least some applications, the individual-based relationships may be more relevant than typical species-level approaches. For example, when parametrizing or testing ecosystem models, trait relationships estimated from individual-level datasets that preserve species differences in abundance and intraspecific variation should more accurately capture ecosystem-level responses compared to analyses based on species means.

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 means or environmental conditions in isolation and allows for robust predictions for species and regions not included in the training data. Across the eastern USA, interspecific trait variation, geographic variation in species' abundances, and intraspecific variation all have important effects on trait distributions

and on inter-trait relationships. 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 vegetation inventory data become available globally, combined modeling approaches, such as the one presented here, can be applied to other geographic regions to study intraspecific variation and trait-species-environment relationships. Such analyses are already possible using national forest inventories in some regions (e.g., some European countries, Canada, New Zealand; Schelhaas et al, 2006, Rati et al., 2018, Gills et al., 2005, Paul et al., 2021) in combination with ever-growing plant trait datasets (e.g., TRY; https://www.try-db.org/). Expanding this work outside the USA could contribute to further understanding the mechanisms driving trait distributions across scales and the link between traits, species distributions, community assembly, and ecosystem function.

Acknowledgements

This work was supported by the Gordon and Betty Moore Foundation's Data-Driven Discovery Initiative through grant GBMF4563 to E. P. White; National Science Foundation (NSF) grant 1926542 to E. P. White, S. A. Bohlman, A. Zare, D. Z. Wang, and A. Singh; NSF grant DEB-1442280 to J. W. Lichstein and S. A. Bohlman; USDA/NIFA McIntire-Stennis program grant FLA-FOR-005470 to S. A. Bohlman; and a University of Florida Biodiversity Institute (UFBI) and

Informatics Institute (UFII) Graduate Fellowship to Sergio Marconi. There was no additional external funding received for this study.

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.

Data accessibility: All data and data products are archived and accessible on Zenodo (https://doi.org/10.5281/zenodo.4651366). Code to replicate the analyses and retrieve data is stored on GitHub (https://github.com/MarconiS/Disentangling-the-role-of-phylogeny-and-climate-on-joint-leaf-trait-distributions-across-Eastern-Uni/tree/0.1).

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. Bechtold, W. A., and P. L. Patterson. 2005. The enhanced Forest Inventory and Analysis program - National sampling design and estimation procedures. General Technical Report SRS-80. USDA Forest Service, Southern Research Station, Asheville, NC.
4. Bonan, G.B., Levis, S., Kergoat, L. & Oleson, K.W. (2002). Landscapes as patches of plant functional types: an integrating concept for climate and ecosystem models. Global Biogeochem. Cy., 16, 1021.
5. Bruelheide, H., Dengler, J., Purschke, O., Lenoir, J., Jiménez-Alfaro, B., Hennekens, S.M., Botta-Dukát, Z., Chytrý, M., Field, R., Jansen, F. and Kattge, J., 2018. Global trait–environment relationships of plant communities. Nature ecology & evolution, 2(12), p.1906.
6. Bürkner 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
7. 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.
8. 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.
9. 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.

10. Chave, J., 2013. The problem of pattern and scale in ecology: what have we learned in 20 years?. Ecology letters, 16, pp.4-16.
11. Clark, J. S. (2016). Why species tell more about traits than traits about species: predictive analysis. Ecology, 97(8), 1979-1993.
12. 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
13. 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.
14. de Bello, F., et al., 2011. Quantifying the relevance of intraspecific trait variability for functional diversity. - Methods in Ecology and Evolution 2: 163-174.
15. 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.
16. Dantas de Paula, M., M. Forrest, L. Langan, J. Bendix, J. Homeier, A. Velescu, W. Wilcke, and T. Hickler. 2021. Nutrient cycling drives plant community trait assembly and ecosystem functioning in a tropical mountain biodiversity hotspot. New Phytologist 232:551–566.

17. Diaz, S., and M. Cabido. 2001. Vive la difference: plant functional diversity matters to ecosystem processes. Trends in Ecology & Evolution 16:646–655.
18. 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.
19. 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.
20. 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.
21. 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.
22. 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.
23. 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.

24. 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.
25. Hart, S. P., et al., 2016. How variation between individuals affects species coexistence. - Ecology Letters 19: 825-838.
26. Hedin, L.O., 2004. Global organization of terrestrial plant–nutrient interactions. Proceedings of the National Academy of Sciences, 101(30), pp.10849-10850.
27. Henn, J.J., Buzzard, V., Enquist, B.J., Halbritter, A.H., Klanderud, K., Maitner, B.S., Michaletz, S.T., Pötsch, 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.
28. 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
29. Kattge, J., Bönisch, G., Díaz, 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.
30. 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.
31. 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.

32. Laughlin, D. C., et al., 2012. A predictive model of community assembly that incorporates intraspecific trait variation. - Ecology Letters 15: 1291-1299.
33. Lepš, J., et al., 2011. Community trait response to environment: disentangling species turnover vs intraspecific trait variability effects. - Ecography 34: 856-863.
34. Lima, Taysla RA, et al., "Lignin composition is related to xylem embolism resistance and leaf life span in trees in a tropical semiarid climate." New Phytologist 219.4 (2018): 1252-1262.
35. Lusk, C.H., Reich, P.B., Montgomery, R.A., Ackerly, D.D. & Cavender-Bares, J. (2008). Why are evergreen leaves so contrary about shade? Trends Ecol. Evol., 23, 299–303.
36. 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.
37. 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
38. 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.
39. McIntyre, S., Díaz, S., Lavorel, S., & Cramer, W. (1999). Plant functional types and disturbance dynamics–Introduction. Journal of Vegetation Science, 10(5), 603-608.
40. 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.

41. 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.
42. 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.
43. Münzbergová, Z., Hadincová, V., Skálová, 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.
44. Muñoz A.-R. & Real R. (2006) Assessing the potential range expansion of the exotic monk parakeet in Spain. Diversity and Distributions 12: 656-665.
45. 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
46. Niinemets, Ü., 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.

47. 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.
48. Ordoñez, 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.
49. 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.
50. 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).
51. Peaucelle, M., Bacour, C., Ciais, P., Vuichard, N., Kuppel, S., Peñuelas, 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.
52. Pecl, G.T., Araújo, M.B., Bell, J.D., Blanchard, J., Bonebrake, T.C., Chen, I.C., Clark, T.D., Colwell, R.K., Danielsen, F., Evengård, B. and Falconi, L., 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science, 355(6332), p.eaai9214.
53. Pérez-Harguindeguy, N., et al., 2013. New handbook for standardised measurement of plant functional traits worldwide. - Australian Journal of Botany 61: 167-234.

54. Poorter, H., Niinemets, Ü., 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.

55. Räty, 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.

56. 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

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

58. 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

59. 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.

60. 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.

61. 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.

62. Scheiter, S., L. Langan, and S. I. Higgins. 2013. Next-generation dynamic global vegetation models: learning from community ecology. New Phytologist 198:957–969.
63. 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/
64. 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.
65. 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.
66. Smith, W.B., 2002. Forest inventory and analysis: a national inventory and monitoring program. Environmental pollution, 116, pp.S233-S242.
67. 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.
68. 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.
69. Swenson, N.G., Weiser, M.D., Mao, L., Araújo, M.B., Diniz-Filho, J.A.F., Kollmann, J., Nogués-Bravo, D., Normand, S., Rodríguez, M.A., García-Valdés, 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.

70. Swenson, N.G., 2014. Phylogenetic imputation of plant functional trait databases. Ecography, 37(2), pp.105-110.
71. 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.
72. 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.
73. Tautenhahn, S., Migliavacca, M. & Kattge, J. (2020). News on intra-specific trait variation, species sorting, and optimality theory for functional biogeography and beyond. New Phytol., 228, 6–10.
74. 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.
75. 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
76. 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.

77. Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M.B., Balaguer, L., Benito-Garzón, 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.

78. Verheijen, L.M., Brovkin, V., Aerts, R., Boenisch, G., Cornelissen, J.H.C., Kattge, J., et al., (2013). Impacts of trait variation through observed trait-climate relationships on performance of an Earth system model: a conceptual analysis. Biogeosciences, 10, 5497–5515.

79. 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.

80. 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.

81. 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.

82. 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.

83. 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.

84. 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.

85. Wullschleger, S. D., H. E. Epstein, E. O. Box, E. S. Euskirchen, S. Goswami, C. M. Iversen, J. Kattge, R. J. Norby, P. M. van Bodegom, and X. Xu. 2014. Plant functional types in Earth system models: past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. Annals of Botany 114:1–16.

86. 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

87. 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.

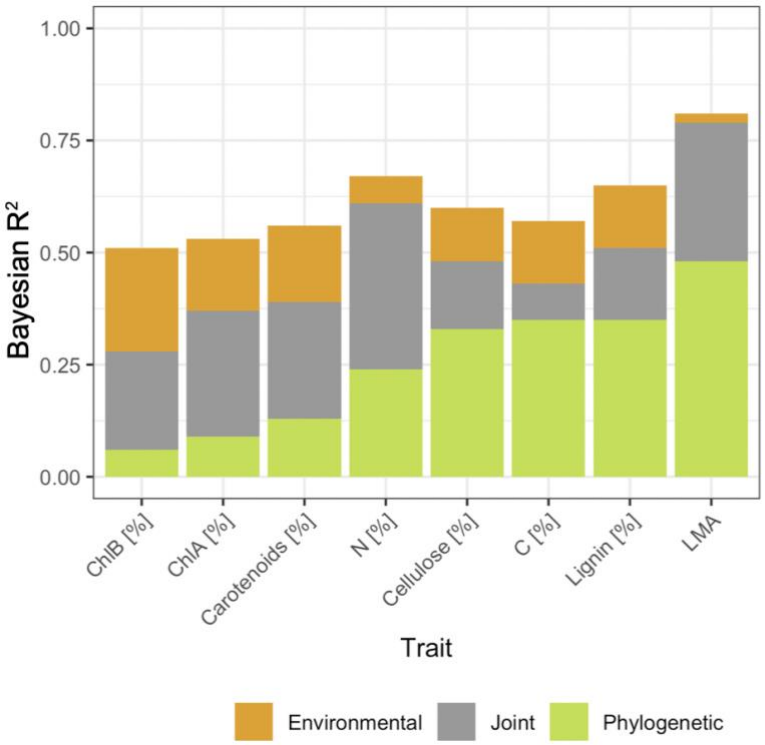

Figure 1. Variance partitioning between pure species effects, pure environmental effects (intraspecific variation), and joint species-environment effects based on the R^2 for independent validation data. Pure environmental variation is independent of species variation, and vice versa. Joint variation is jointly explained by species and environmental effects and cannot be uniquely ascribed to either. LMA is leaf dry mass per area, and all other traits are % mass.

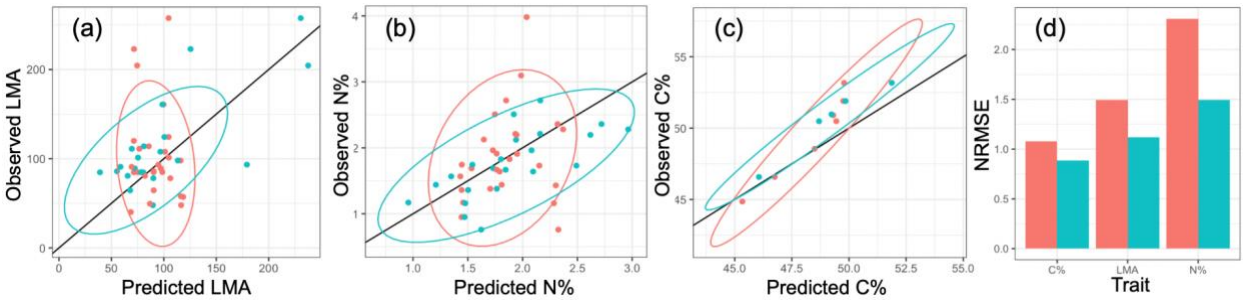

Figure 2. Observed vs. predicted values from the combined model for (a) LMA (g m-2), (b) N%, and (c) C%, either assuming phylogenetically-structured covariance among species random effects (azure), or assuming independent species random effects (orange). Ellipses are 95% confidence ellipses. (d) RMSE (normalized by standard deviation) showing the difference between accounting for (azure) or ignoring (orange) phylogenetic relationships when predicting traits for species not included in the training data.

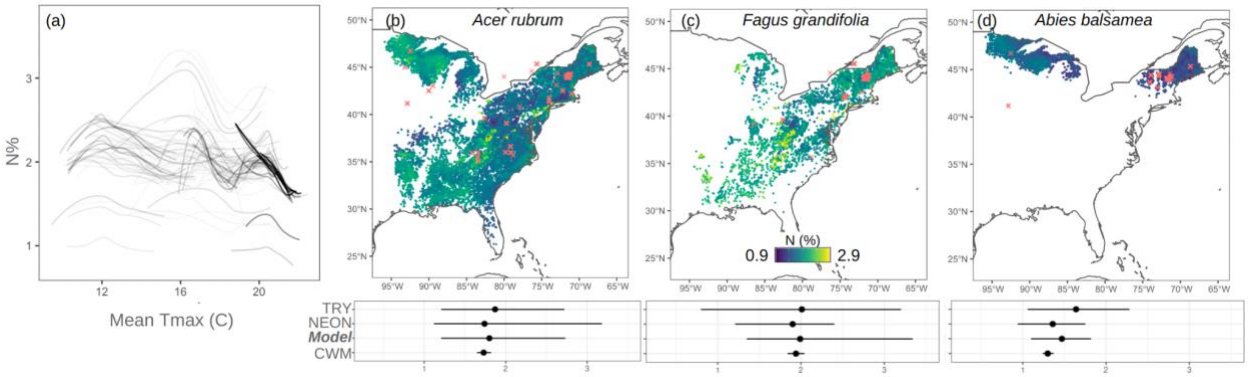

Figure 3. Intraspecific variation of predicted leaf N% for tree species in the eastern USA. (a) Relationships (estimated using GAMs) of N% (median predicted values) vs. maximum monthly temperature for the 91 species (each species has a separate line) showing significant effects of temperature on N% (the darkness of the curve increases with the strength of the relationship between N% and temperature). (b-d) Geographic patterns and range of intraspecific N% for two widespread deciduous broadleaf (*Acer rubrum* and *Fagus grandifolia*) and one widespread needleleaf evergreen (*Abies balsamea*) tree species. Mean (points) and 99% intraspecific ranges (lines) in b-d are from: TRY data (within the eastern USA), NEON data (within the eastern USA), the combined model, and random error from the species-only model. Red crosses represent TRY sample locations.

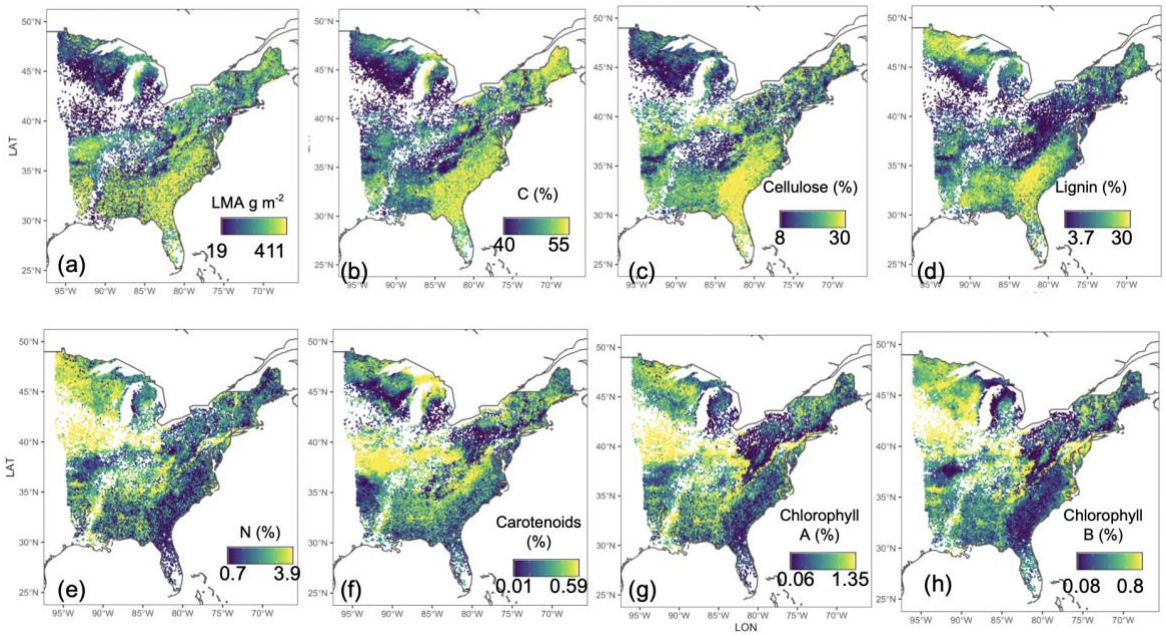

Figure 4. Distributions across the eastern USA for leaf structural (LMA, lignin, cellulose and C%), macro-nutrient (N%), and pigment (carotenoids, chlorophyll A and B) traits. Values represent trait averages at each FIA plot (n = 30,331 plots), calculated across all individual trees sampled within the plot.

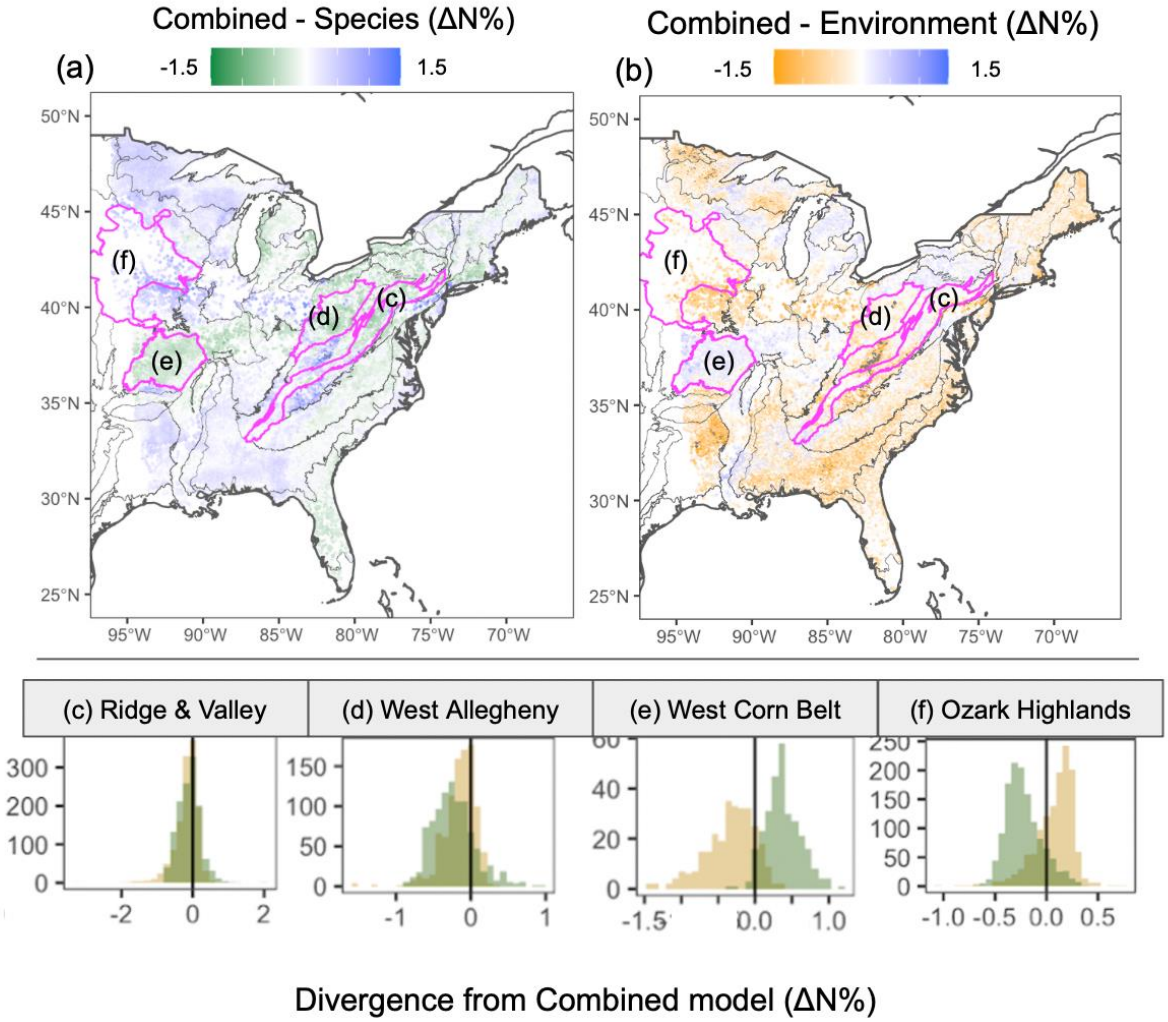

Figure 5. Divergence in N% between the combined model and the species-only (a) and environment-only (b) models. Dots represent FIA plots where predicted N% was higher in the combined (blue), environment (orange), or species (green) model. White areas represent plots with little to no divergence ($p > 0.0001$), or areas with no FIA plots recorded. (c-f) Examples of divergence patterns in four ecoregions (see the outlined ecoregions in the maps, labeled c-f according to their corresponding histogram): combined model minus environment model (orange) and combined model minus species model (green). Vertical black lines represent zero divergence. Ecoregion examples include: all three models capturing similar information with ΔN% near 0 (c); species and environment models similar to each other but diverging from the combined model (d); and and divergence among all three models (e-f).

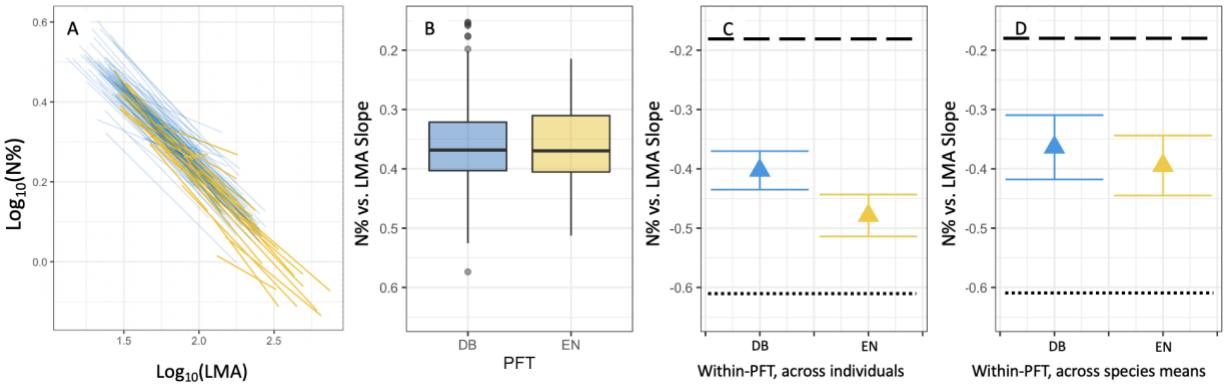

Figure 6. Relationships between leaf nitrogen percent mass (N%) and leaf mass per area (LMA) across the eastern USA, based on combined-model predictions of N% and LMA for individual trees in eastern USA Forest Inventory and Analysis (FIA) plots. (A) Fitted intraspecific ordinary least squares (OLS) regression lines for eastern USA deciduous broadleaf (DB: blue) and evergreen needleleaf (EN: orange) species. Each line is an individual species. (B) Distribution of intraspecific OLS slopes for DB and EN species; bars are medians, boxes span the interquartile range, whiskers span 95% of slopes, and points show outliers. (C) OLS slopes (± s.e.) across all individuals within each PFT. (D) OLS slopes (± s.e.) across species means within each PFT. Individual trait values were $\log_{10}$-transformed for analyses in (A-C), and species means were $\log_{10}$-transformed for analyses in (D). Dashed and dotted black lines in panels C and D are from Osnas et al., (2018): the dashed line at -0.18 represents the intraspecific slope across the canopy light gradient, and the dotted black line at -0.62 represents the slope across global species (see details in Supplement 6).