Mapping local and global variability in plant trait distributions

Ethan E. Butler[a,1,2], Abhirup Datta[b,1,2], Habacuc Flores-Moreno[a,c], Ming Chen[a], Kirk R. Wythers[a], Farideh Fazayeli[d], Arindam Banerjee[d], Owen K. Atkin[e,f], Jens Kattge[g,h], Bernard Amiaud[i], Benjamin Blonder[j], Gerhard Boenisch[g], Ben Bond-Lamberty[k], Kerry A. Brown[l], Chaeho Byun[m], Giandiego Campetella[n], Bruno E. L. Cerabolini[o], Johannes H. C. Cornelissen[p], Joseph M. Craine[q], Dylan Craven[h,r], Franciska T. de Vries[s], Sandra Díaz[t,u], Tomas F. Domingues[v], Estelle Forey[w], Andrés González-Melo[x], Nicolas Gross[y,z,aa], Wenxuan Han[bb,cc], Wesley N. Hattingh[dd], Thomas Hickler[ee,ff], Steven Jansen[gg], Koen Kramer[hh,ii], Nathan J. B. Kraft[jj], Hiroko Kurokawa[kk], Daniel C. Laughlin[ll], Patrick Meir[f,mm], Vanessa Minden[nn], Ülo Niinemets[oo], Yusuke Onoda[pp], Josep Peñuelas[qq,rr], Quentin Read[ss], Lawren Sack[jj], Brandon Schamp[tt], Nadejda A. Soudzilovskaia[uu], Marko J. Spasojevic[vv], Enio Sosinski[ww], Peter E. Thornton[xx], Fernando Valladares[yy], Peter M. van Bodegom[uu], Mathew Williams[mm], Christian Wirth[g,h,zz], and Peter B. Reich[a,aaa]

Edited by William H. Schlesinger, Cary Institute of Ecosystem Studies, Millbrook, NY, and approved October 18, 2017 (received for review May 31, 2017)

Our ability to understand and predict the response of ecosystems to a changing environment depends on quantifying vegetation functional diversity. However, representing this diversity at the global scale is challenging. Typically, in Earth system models, characterization of plant diversity has been limited to grouping related species into plant functional types (PFTs), with all trait variation in a PFT collapsed into a single mean value that is applied globally. Using the largest global plant trait database and state of the art Bayesian modeling, we created fine-grained global maps of plant trait distributions that can be applied to Earth system models. Focusing on a set of plant traits closely coupled to photosynthesis and foliar respiration—specific leaf area (SLA) and dry mass-based concentrations of leaf nitrogen (N_m) and phosphorus (P_m), we characterize how traits vary within and among over 50,000 ~50 × 50-km cells across the entire vegetated land surface. We do this in several ways—without defining the PFT of each grid cell and using 4 or 14 PFTs; each model's predictions are evaluated against out-of-sample data. This endeavor advances prior trait mapping by generating global maps that preserve variability across scales by using modern Bayesian spatial statistical modeling in combination with a database over three times larger than that in previous analyses. Our maps reveal that the most diverse grid cells possess trait variability close to the range of global PFT means.

plant traits | Bayesian modeling | spatial statistics | global | climate

Modeling global climate and the carbon cycle with Earth system models (ESMs) requires maps of plant traits that play key roles in leaf- and ecosystem-level metabolic processes (1–4). Multiple traits are critical to both photosynthesis and respiration, foremost leaf nitrogen concentration (N_m) and specific leaf area (SLA) (5–7). More recently, variation in leaf phosphorus concentration (P_m) has also been linked to variation in photosynthesis and foliar respiration (7–12). Estimating detailed global geographic patterns of these traits and corresponding trait–environment relationships has been hampered by limited measurements (13), but recent improvements in data coverage (14) allow for greater detail in spatial estimates of these key traits.

Previous work has extrapolated trait measurements across continental or larger regions through three methodologies: (*i*) grouping measurements of individuals into larger categories that share a set of properties [a working definition of plant functional types (PFTs)] (4, 15), (*ii*) exploiting trait–environment relationships (e.g., leaf N_m and mean annual temperature) (1, 16–20), or (*iii*) restricting the analysis to species whose presence has been widely estimated on the ground (21–24). Each of these methods has limitations—for example, trait–environment relationships do not well explain observed trait spatial patterns (1, 25), while species-based approaches limit the scope of extrapolation to only areas with well-measured species abundance. More critically, the first two global methodologies emphasized estimating a single trait value per PFT at every location, whereas both ground-based (5, 14) and remotely sensed (26) observations suggest that at ecosystem or landscape scales traits would be better represented by distributions. Here, we use an updated version of the largest global database of plant traits (14) coupled with modern Bayesian spatial statistical modeling techniques (27) to capture local and global variability in plant traits. This combination allows the representation of trait variation both within pixels on a gridded land surface and across global environmental gradients.

Information is lost when the range of measured trait values is compressed into a single PFT (Fig. 1). We observe that the global range of site-level SLA values for a single PFT such as broadleaf evergreen tropical trees (Fig. 1 *A* and *C*) is quite large (2.7–65.2 $m^2 \cdot kg^{-1}$). Even after limiting the scope to a single

Significance

Currently, Earth system models (ESMs) represent variation in plant life through the presence of a small set of plant functional types (PFTs), each of which accounts for hundreds or thousands of species across thousands of vegetated grid cells on land. By expanding plant traits from a single mean value per PFT to a full distribution per PFT that varies among grid cells, the trait variation present in nature is restored and may be propagated to estimates of ecosystem processes. Indeed, critical ecosystem processes tend to depend on the full trait distribution, which therefore needs to be represented accurately. These maps reintroduce substantial local variation and will allow for a more accurate representation of the land surface in ESMs.

Author contributions: E.E.B., A.D., H.F.-M., M.C., K.R.W., F.F., A.B., O.K.A., J.K., and P.B.R. designed research; E.E.B. and A.D. performed research; E.E.B., A.D., H.F.-M., and J.K. analyzed data; and E.E.B., A.D., H.F.-M., M.C., K.R.W., A.B., O.K.A., J.K., B.A., B.B., G.B., B.B.-L., K.A.B., C.B., G.C., B.E.L.C., J.H.C.C., J.M.C., D.C., F.T.d.V., S.D., T.F.D., E.F., A.G.-M., N.G., W.H., W.N.H., T.H., S.J., K.K., N.J.B.K., H.K., D.C.L., P.M., V.M., Ü.N., Y.O., J.P., Q.R., L.S., B.S., N.A.S., M.J.S., E.S., P.E.T., F.V., P.M.v.B., M.W., C.W., and P.B.R. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The code and data necessary to run the models are available at https://github.com/abhirupdatta/global_maps_of_plant_traits.

[1]E.E.B. and A.D. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: eebutler@umn.edu or abhidatta@jhu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1708984114/-/DCSupplemental.

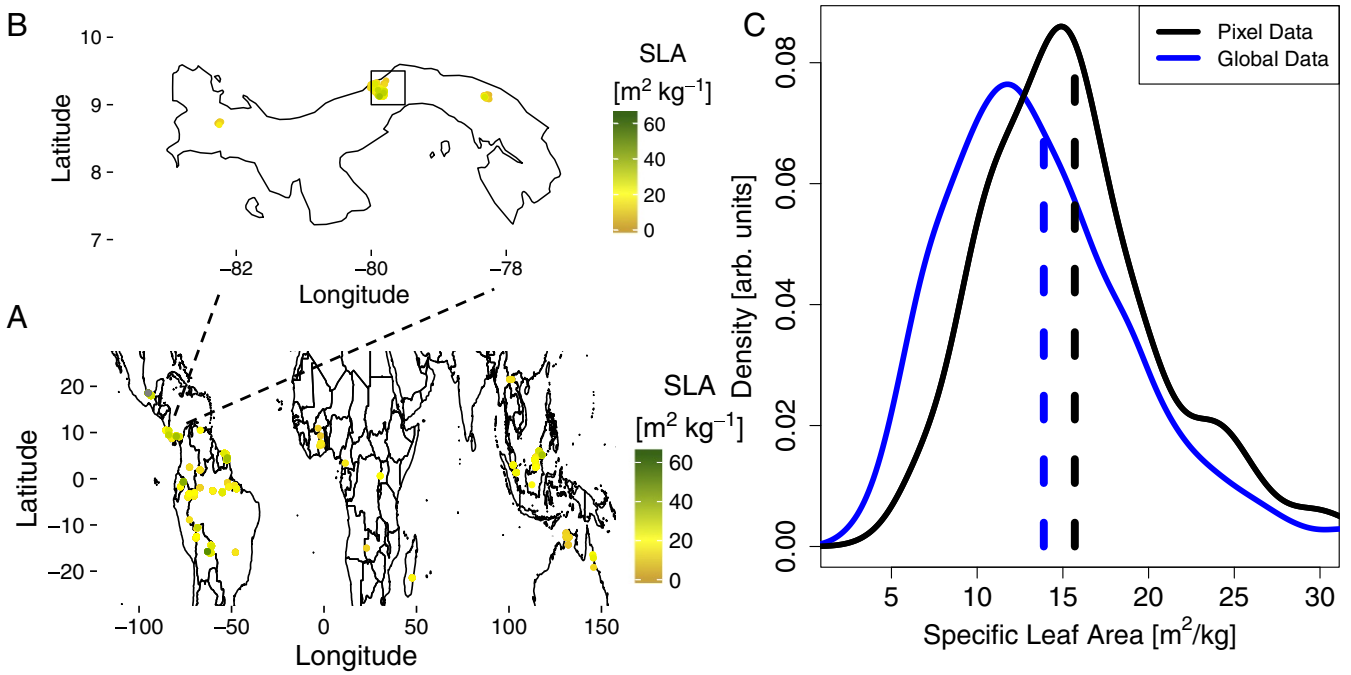

Fig. 1. Trait data. (*A*) Global locations and values of specific leaf area measurements for the PFT tropical broadleaf evergreen trees. (*B*) Locations and values of specific leaf area measurements for the tropical broadleaf evergreen trees in Panama. The square in the center indicates a 0.5° × 0.5° pixel containing the Barro Colorado Island sites (see Fig. 5). These points have been jittered up to 0.05° to highlight the density of measurements. (*C*) The full distribution of specific leaf area values for all species classified as evergreen broadleaf tropical trees. The blue line is the global data while black is the local pixel, and the dashed vertical lines are the respective means.

well-measured 0.5° × 0.5° pixel within Panama (Fig. 1 *B* and *C*), there is still a wide range of SLA values (4.7–37.7 $m^2 \cdot kg^{-1}$) with a local mean of 15.7 $m^2 \cdot kg^{-1}$ and a local standard deviation of 5.4 $m^2 \cdot kg^{-1}$—over one-third of the local mean. By contrast, the mean SLA value of all species associated with broadleaf evergreen tropical trees is 13.9 $m^2 \cdot kg^{-1}$, over 10% lower than the local average (Fig. 1*C*). Thus, single trait values per PFT fail to capture variability in trait values within or among grid cells, i.e., over a wide range of spatial scales.

Transitioning from a single trait value per PFT (within or among grid cells) to a distribution may lead to significantly different modeling results (20) as critical plant processes, such as photosynthesis, are nonlinear with respect to these traits (28). This is reinforced by recent modeling studies that have begun to incorporate distributions of traits at regional (29, 30) and global (31) scales. It has been shown that using trait distributions leads to different estimates of carbon dynamics (32) and that higher-order moments of trait distributions contribute to sustaining multiple ecosystem functions (33). While species-level mapping (21, 23, 24) does capture trait distributions, it has been limited geographically and restricted to subsets of functional groups.

Even the largest plant trait database offers only partial coverage across the globe in terms of site-level measurements. Hence, gap-filling approaches need to be adopted to extrapolate trait values at regions with no data coverage. Here, we overcome data limitations through PFT classification, trait–environment relationships, and additional location information to develop a suite of models capable of estimating trait distributions across the entire vegetated globe. The simplest one is a categorical model, which assigns traits to maps of remotely sensed PFTs. Every species, with its corresponding trait values, is associated with a PFT and these trait distributions are extrapolated to the satellite-estimated range of the PFT (*SI Appendix*, Figs. S1 and S2). The second one is a Bayesian linear model that complements the PFT information with trait–environment relationships. The third one is a Bayesian spatial model that, in addition to PFTs and the trait–environment relationships, leverages additional location information via Gaussian processes (*Materials and Methods*). The use of a spatial Gaussian process in this context is unique and model evaluation reveals the superior predictive performance of this model.

Each of these methods interpolates (and extrapolates) both mean trait values and entire trait distributions across space (i.e., across grid cells on a global map). These models are further stratified by three different levels of PFT categorization: (*i*) PFT-free, all plants in a single group (i.e., no PFTs); (*ii*) broad, 4 groups based on growth form and leaf type; and (*iii*) narrow, 14 groups based on further environmental, phenological, and photosynthetic categories (*Materials and Methods*). The PFT-free categorization groups all plants into a single class, while the broad grouping (4-PFT) is similar to the vegetation classification used in the Joint UK Land Environment Simulator land surface (34), and the narrow (14-PFT) category is equivalent to the classification used in the Community Land Model (CLM) (4, 15, 35).

The abovementioned methods allow for a representation of global vegetation that enables a more accurate formulation of functional diversity than the single-trait value per PFT paradigm that is widely used (4). The traits studied here—SLA, N_m, and P_m—are central to predicting variation in rates of plant photosynthesis (5, 6, 9, 11) and foliar respiration (10, 36). The importance of these traits and the more advanced representation of functional diversity developed here may be used to better capture the response of the land surface component of the Earth system to environmental change.

Results and Discussion

Model Evaluation. Given the full suite of nine models proposed, we conducted extensive model evaluation (Table 1) to determine the trade-offs associated with each methodology and resolution

Table 1. Model evaluation

Model	ps-R^2, %	RMSPE	CP, %
SLA			
Cf	NA	8.13	91.2
Cb	16.9	7.13	94.7
Cn	26.0	6.66	95.8
Lf	4.6	7.99	91.3
Lb	23.4	6.93	94.0
Ln	30.7	6.53	**95.2**
Sf	45.5	7.54	93.6
Sb	58.5	6.31	97.7
Sn	**60.2**	**6.13**	97.7
N_m			
Cf	NA	7.16	93.3
Cb	12.5	6.95	93.2
Cn	19.4	6.47	92.7
Lf	5.2	7.28	93.2
Lb	16.7	6.71	94.3
Ln	24.1	6.42	**94.6**
Sf	44.2	7.19	93.6
Sb	53.7	6.36	96.1
Sn	**54.8**	**6.18**	96.1
P_m			
Cf	NA	0.86	90.5
Cb	5.3	0.86	90.5
Cn	28.1	**0.78**	91.1
Lf	25.6	0.84	87.2
Lb	32.8	0.85	85.3
Ln	35.4	0.82	87.0
Sf	62.0	0.83	90.7
Sb	66.7	0.81	**92.0**
Sn	**67.6**	0.80	91.3

Shown are the pseudo-R^2 (ps-R^2), RMSPE, and CP statistics for all nine models, for each of the three traits. The entries in boldface type correspond to the model producing highest ps-R^2, lowest RMSPE, or CP closest to 0.95. The categorical PFT-free model (Cf) produces a constant estimate and hence ps-R^2 is not defined. Each model is indicated by a two-letter abbreviation: C, categorical (no regression); L, linear (linear regression); and S, spatial (linear regression with spatial term) and the accompanying PFT resolution: f, PFT-free (no PFT information); b, broad (4-PFT); and n, narrow (14-PFT).

of PFT. We assessed the predictive capability of the models, using the root-mean-square predictive error (RMSPE) based on out-of-sample data (*SI Appendix*, section S6). Among the nine models, the spatial narrow 14-PFT model emerged as the best predictor of mean trait values for SLA and N_m and the second best for P_m (Table 1). However, the spatial broad 4-PFT model performed nearly as well (Table 1). The models' abilities to correctly estimate the spread of the trait distributions were assessed using the out-of-sample coverage probabilities (CPs)—the proportion of instances the model-predicted 95% confidence intervals contained the observed trait values. Most of the models provided adequate coverage (CP of around 90% or more). See *SI Appendix*, section S4, for more detailed definitions of the model comparison metrics.

The improvement in prediction afforded by the inclusion of (*i*) a spatial term and (*ii*) PFT information (Table 1) invites further examination. First, the spatial term in our model likely incorporates some of the finer-scale variation that is unavailable given the relatively large grid cell size of the environmental covariates used in global studies. Thus, the spatial term allows for adjustment of trait values among neighboring or regional grid cells that the relatively coarse environmental metrics are not able to capture. Finer-scale studies that can evaluate local variations in climate, soil, or other relevant abiotic or biotic covariates may see less improvement from the inclusion of a spatial term, as they may directly measure local sources of variation. Second, the use of PFTs greatly improves the models, perhaps for similar reasons involving the degree of variation the raw data fail to incorporate. The greatest decrease in RMSPE occurs between the PFT-free grouping (a single category for all plants) and the broad (4-PFT) grouping across each of the models tested. If our trait data were perfectly predicted by environment, there would be no usefulness to including PFTs in mapping traits. That this is not so implies that the broad PFTs, based primarily on growth form and leaf type, offer superior predictive skill than environmental covariates on their own (19). However, the extra information in the narrow (14-PFT) grouping does further improve the fit and produces the most accurate predicted trait surface.

Global Maps. We selected two sets of maps to describe, in broad strokes, how trait distributions vary across the land surface: the narrow 14-PFT spatial model and its categorical counterpart. The narrow 14-PFT spatial model is the best predictor of mean trait values and provided adequate coverage probability (Figs. 2 *A* and *B*, 3 *A* and *B*, and 4 *A* and *B*). For comparison, we also include the 14-PFT categorical model, which is most similar to maps currently used in ESMs (Figs. 2 *C* and *D*, 3 *C* and *D*, and 4 *C* and *D*). Maps for the other models can be found in *SI Appendix*, Figs. S8–S16. The mean and SD are presented as a summary of the full log-normal distribution within each pixel, but there are full distributions estimated in each pixel (*Case Studies*).

The SD maps (Figs. 2 *B* and *D*, 3 *B* and *D*, and 4 *B* and *D*) compared with the mean maps (Figs. 2 *A* and *C*, 3 *A* and *C*, and 4 *A* and *C*) highlight one of the central results of this analysis—the local SDs of trait values are of similar magnitudes to their respective means. Generally, we observed that the local SD is close to half the local mean value but can approach the global range of the trait mean values; e.g., N_m (Fig. 3) has a maximum local SD of 9 mg N/g, and the global mean range is only ∼10 mg N/g. The maps of the trait SDs follow similar patterns to the means, although there are several regions where the mean varies more markedly than the SD, such as SLA in the southeast United States and China in the categorical model (Fig. 2 *C* and *D*) and similarly for N_m in the spatial model across the Sahel in sub-Saharan Africa (Fig. 3 *A* and *C*). The lack of variation in the SD is most clear in the categorical model for N_m while both models show relatively modest variation in P_m.

For each of the three traits, the broad features of both the categorical and spatial models are similar, but there are numerous marked differences across regional and fine spatial scales (Figs. 2–4). The shared broad features of the maps from both models include SLA (Fig. 2) and P_m (Fig. 4) increasing from the tropics to the poles, while N_m (Fig. 3) has more modest variation, except that it tends to be lower in regions dominated by needle-leaved trees. Some of the notable differences between the models include the spatial model's greater range and more marked variability of SLA within equatorial regimes (e.g., Brazil or central Africa); it also captures the low SLA of most of arid Australia better than the categorical model (Fig. 2*A*) and more strongly highlights the gradient of P_m from the tropics to the Arctic (16) (Fig. 4*A*).

The most consistent estimates between the categorical and spatial models are in the boreal regions dominated by needle-leaved trees; the measurements in this region are relatively sparse, which may have limited the ability of the spatial model to capture differences. On the other hand, broad-leaved trees span a wide range of environments, but a large portion of the measurements come from the tropics (66%), where there is a limited range of values among the climate covariates and therefore little variation with which to estimate a correlation. The

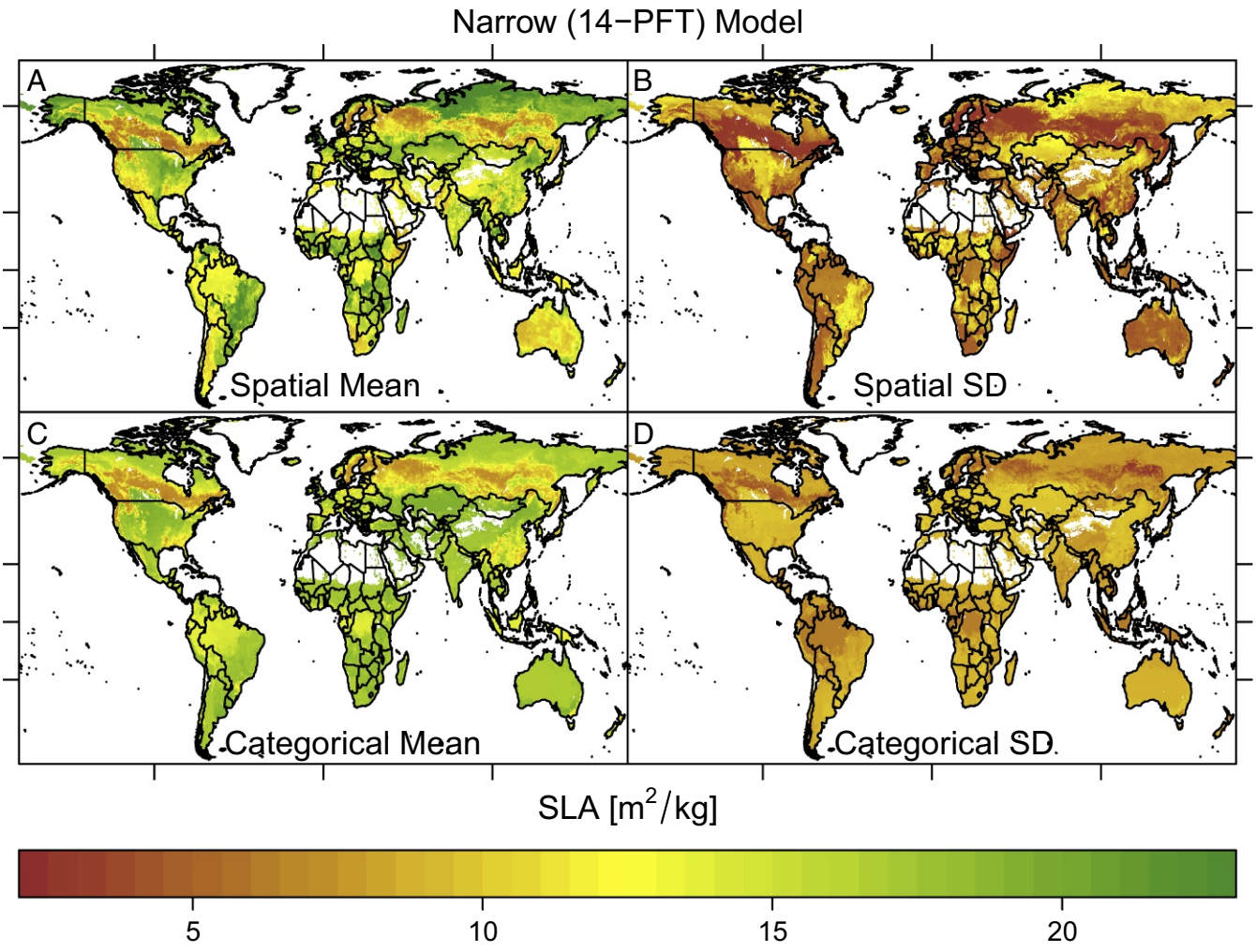

Fig. 2. SLA maps. (*A* and *B*) Narrow (14-PFT) Bayesian spatial model pixel mean and SD estimates, respectively. (*C* and *D*) Narrow (14-PFT) categorical model pixel mean estimates and SD estimates, respectively. For clarity, the color bars have been truncated at the compound 5th and 95th percentiles of both models. Latitude tick marks indicate the equator, tropics, and Arctic Circle and longitude is marked at 100°W, 0°, and 100°E.

grasses and shrubs have the largest SDs of the four broad PFTs (*SI Appendix*, Table S4) and dominate wide swaths of the land surface, but have fewer measurements—shrubs are the least measured of the broad PFTs in the database, and this appears to reduce the accuracy of the categorical model more than that of the spatial model (Table 1). The fact that shrubs are assumed to dominate in arid and boreal environments, which also tend to be undersampled, also likely contributes to these differences.

Our results also suggest that the breadth of functional niche space is reduced in both boreal and tropical biogeographic regions. The low variation across all three traits within the boreal forest implies that there is strong filtering and smaller niche space available in this relatively harsh environment. Surprisingly, despite the high species diversity in tropical forests, we also find that SLA and P_m have relatively low variation in these forests—suggesting that in this environment the trait space is reduced. This could be, in part, an artifact of the Earth system model PFT classification omitting herbaceous species. Conversely, grasslands and savannahs exhibit large variation in total trait space, suggesting these environments permit a wider range of strategies than in both the boreal and tropical regions. Most broadly, both the data and the spatial model suggest (*SI Appendix*, Figs. S24 and S25) lowest leaf nitrogen values in temperate climates that increase in both cooler and warmer regions; this may indicate a more complicated leaf biochemistry–temperature relationship than has previously been suggested (16).

Case Studies. We conducted two regional case studies to provide a more in-depth analysis of the true and predicted shapes of trait distributions than can be provided by the SD maps and coverage probability. In these case studies trait data were pooled over an area to construct full trait distributions and then formally compared with the model predicted distributions.

We considered two areas with substantially different environmental conditions to evaluate the trait distributions obtained from the spatial and categorical models. We chose a single pixel that contained a highly studied site with numerous measurements of tropical trees, Barro Colorado Island (BCI), Panama; and a collection of pixels in an arid environment in which the mean estimates for SLA of the spatial and categorical models substantially disagreed, the southwestern United States. These areas were in the training data, and this analysis constituted a more detailed analysis of the models' fit to the observed distribution of these locations. Here, the focus was on the structure of the full distribution of traits predicted at these sites; *SI Appendix*, Fig. S17 is a map of the measurements that comprised these locations and other sites included in this analysis. Both areas offer

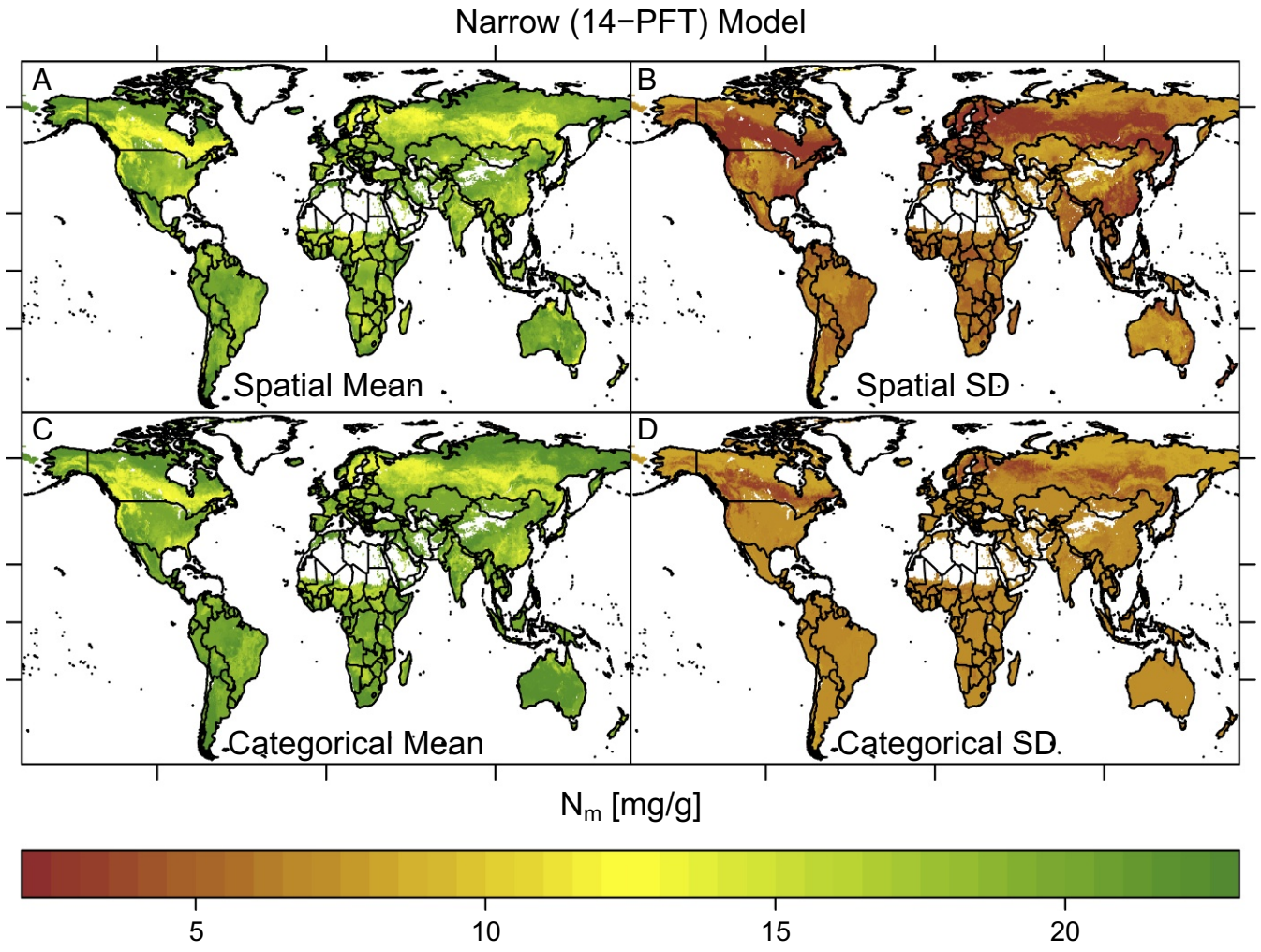

Fig. 3. Nitrogen (mass) maps. (*A* and *B*) Narrow (14-PFT) Bayesian spatial model pixel mean and SD estimates, respectively. (*C* and *D*) Narrow (14-PFT) categorical model pixel mean estimates and SD estimates, respectively. For clarity, the color bars have been truncated at the compound 5th and 95th percentiles of both models. Latitude tick marks indicate the equator, tropics, and Arctic Circle and longitude is marked at 100°W, 0°, and 100°E.

further insight into the structure of the distributions estimated by the categorical and spatial models.

In the pixel containing BCI, the categorical and spatial models broadly agreed for all three traits (Fig. 5 *A*, *C*, and *E*), although the spatial model means were only half as distant from the observed means for SLA and N_m (4% vs. 8% and 5% vs. 10%, respectively). There were only two PFTs present in this pixel: tropical broadleaf evergreen and deciduous trees. Despite the general similarity of the shapes of the distribution, the spatial model appears capable of capturing some subtle features. This is clearest for leaf nitrogen, where the peak of the distribution was quite broad. This is neatly captured in the narrow PFT model, and the pattern was detectable through the Kolmogorov–Smirnov (K-S) statistic, which evaluates the similarity of two full distributions. Indeed, the superiority of the spatial model was reinforced by a closer match for the Bayesian spatial model across all traits at BCI, although for P_m it was the PFT-free spatial model that fitted best (*SI Appendix*, Table S6).

The differences between the trait distributions of the categorical and Bayesian spatial models were stark in the southwestern United States, although the mean estimates for N_m and P_m were close (Fig. 5 *B*, *D*, and *F*). This may be a result of the topographic complexity of this region and the resulting difficulty of aggregating climate and soil covariates at the 0.5° pixel scale and the sparser sampling than at BCI. To get enough data to approximate a distribution, we aggregated 18 pixels with nine PFTs including every temperate category, although many of them are only marginally present. The inclusion of so many PFTs produced a noisier distribution in the categorical model than suggested by the data and estimated by the spatial model. Neither of the models produced distributions that matched as well with the observations; however, it is notable how close the mean values for both models matched the observations for N_m and P_m, and the spatial model did well for the mean SLA.

Environmental Covariates and the Spatial Term. The improvement in prediction from the linear model to the spatial model is partially explained by weak trait–environment relationships (*SI Appendix*, Tables S1–S3). The magnitude of spatial variation explained by the Gaussian process model is comparable to that of the unexplained trait variation. For most of the spatial models, the estimated spatial range was around 300 km; this suggests a strong spatial effect and implies that the spatial model can provide more precise information about the trait distribution near the locations where we have data. This was largely borne out in the case studies and is illustrated more explicitly in Fig. 6

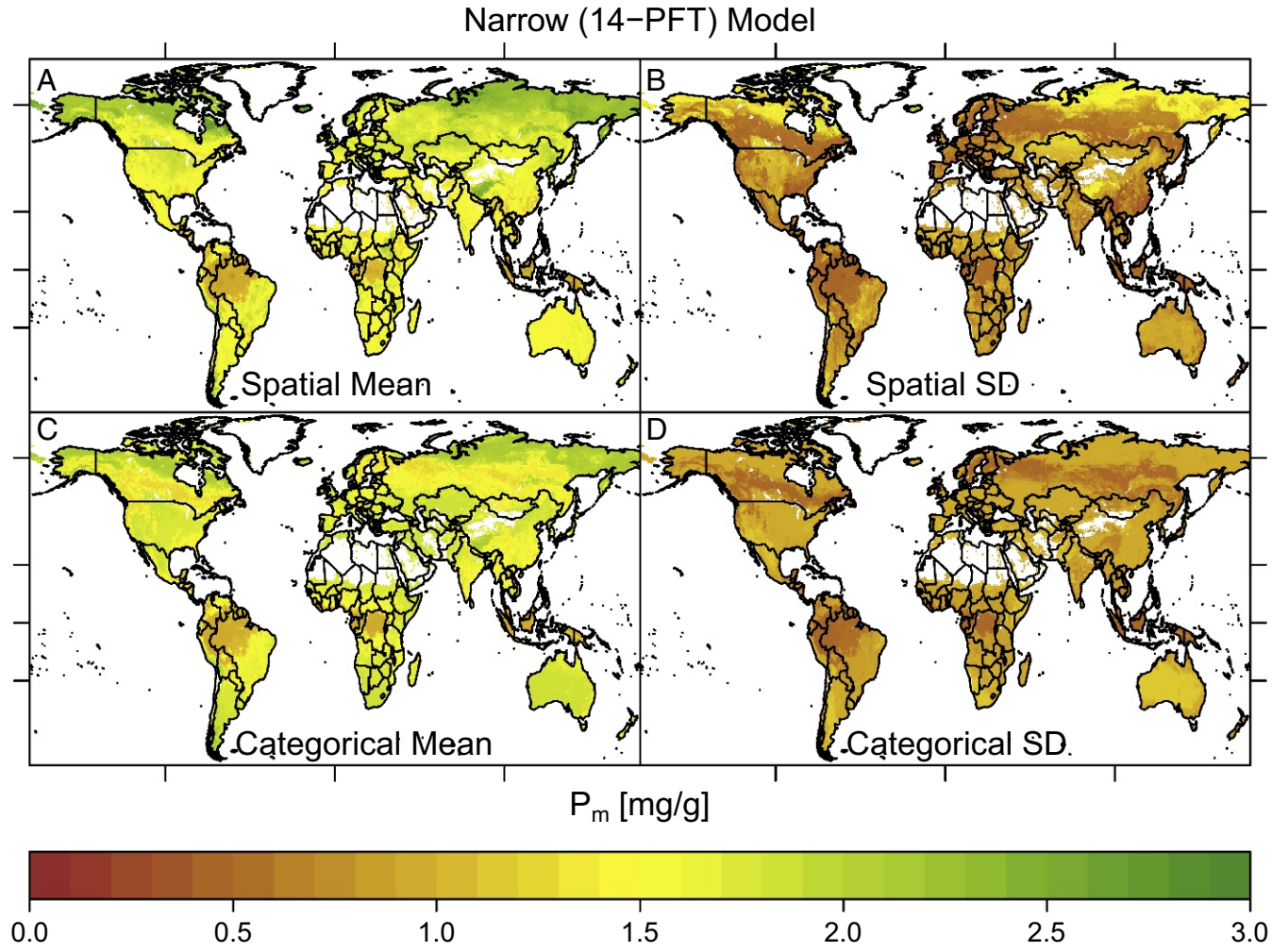

Fig. 4. Phosphorus (mass) maps. (*A* and *B*) Narrow (14-PFT) Bayesian spatial model pixel mean and SD estimates, respectively. (*C* and *D*) Narrow (14-PFT) categorical model pixel mean estimates and SD estimates, respectively. For clarity, the color bars have been truncated at the compound 5th and 95th percentiles of both models. Latitude tick marks indicate the equator, tropics, and Arctic Circle and longitude is marked at 100°W, 0°, and 100°E.

where the predicted trait SD for the spatial model was up to 50% lower than for the linear nonspatial model near locations with trait measurements. The spatial model leverages local information to reduce the uncertainty of trait estimation near data locations and may provide guidance for future data collection by identifying high-uncertainty regions.

Applications for Trait Distributions. Plant traits vary across a range of spatial scales, and the spatial model best captures changes across large spatial gradients (such as in Amazonia and Australia) as well as the subtleties within pixels. Maps for all of the models highlight how much information about local variability is lost when representing plant traits with a single value and suggest that a first application of these maps will be for ESMs to incorporate these scales of variability. For process-based ESMs, the simplest model to incorporate will likely be the categorical model as it is closest to the current PFT approach, but this model is also the least flexible. The more sophisticated models developed here provide more accurate large-scale variation and may be used to infer new trait values in a novel climate by perturbing the climatic covariates (37). However, given the likelihood of nonlinear trait–environment relationships, the spatial sparsity of the data, and the possibility of alternate strategies within a PFT that may alter the trait–environment relationship in a future climate some caution is called for when using these models for extrapolation. Future ecosystem models could also integrate the leaf-level variation in these maps with canopy-scale changes in leaf display traits—leaf angle, azimuth, and total area.

We have emphasized the quality of the Bayesian spatial model with narrow PFTs, but there is an intriguing possibility opened by the PFT-free model (*SI Appendix*, Figs. S8, S11, and S14)—that being the representation of vegetation without reference to PFTs (1). In this case the representation of vegetation would rely entirely on the structure of trait distributions at various landscape scales (1). Such a representation eliminates the need to separately model the future locations of PFTs (or species) when inferring the future distribution of traits; hence, the output of a model like that developed here could be updated with future environmental covariates, with the caveats that "out of sample prediction" may entail. At the same time, this method would allow for greater functional diversity than multiple PFTs with single-trait values, as is currently used in most ESMs. Adopting this approach does, however, raise the issue of how to deal with the paucity of surface observations in some regions, as evidenced by the greater errors associated with estimating out of sample values with this model (Table 1). Complementary work has retrieved leaf trait maps from a global carbon cycle model fused with Earth observations (38), providing another method

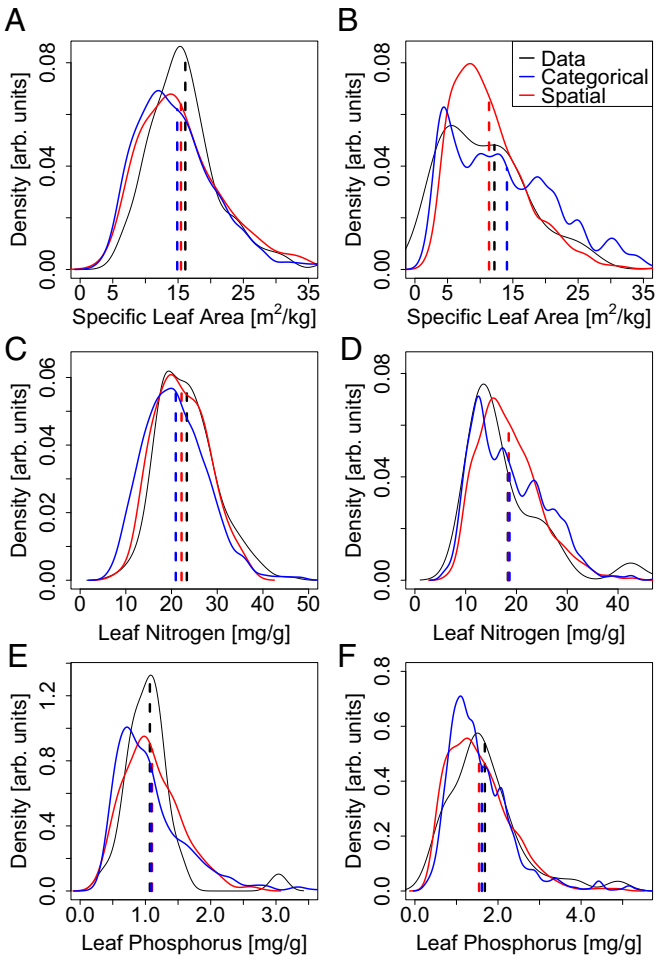

Fig. 5. Empirical trait distributions. Barro Colorado Island (*A*, *C*, and *E*) and the US Southwest (*B*, *D*, and *F*). *A* and *B* show SLA, *C* and *D* show leaf nitrogen, and *E* and *F* show leaf phosphorus. Each panel depicts the distribution of the data in solid black, the categorical model in blue, and the Bayesian spatial model in red. The dashed vertical lines indicate mean values.

that could be used for direct comparison against the trait maps produced here. While the methodology outlined in our analysis brings the possibility of a PFT-free land surface closer, we remain several steps away from being able to make such maps as accurately as we do using PFT characterizations for trait prediction. Several actions can bring us closer to that goal. First, incorporation of additional information (such as phylogenetic relatedness and trait–trait covariance) will likely improve trait maps, even using existing observations. Second, as the current level of observations is extremely sparse in some regions and sparse in most regions, expanded trait databases will also aid in development of PFT-free trait maps.

Conclusions

SLA and N_m are essential inputs into the land surface components of Earth system models, and while phosphorus has not yet been as widely incorporated into ESMs, it has been shown—particularly across the tropics—to be important to photosynthesis (9, 11, 39–42) and respiration (11, 12, 36). The maps and trait–environment relationships presented here may be used by existing land surface models that use similar categories to classify vegetation. However, it should be noted that PFT-dependent models often have many other parameters that have been calibrated to historical estimates of particular trait values (4). Thus, the values developed here, while likely drawing from a larger pool of measurements than has been done previously, cannot necessarily be adopted without further modification of other model elements (37, 43). Nonetheless, these results can be incorporated into a wide class of models with relative ease. We can now provide global trait distributions at the pixel scale.

The global land surface is perhaps the most heterogeneous component of the Earth system. Reducing vegetation to a collection of PFTs with fixed trait values has been the preferred method to constrain this heterogeneity and group similar biochemical and biophysical properties; however, this has been at the expense of functional diversity. This analysis quantifies the substantial magnitude of this ignored trait variation. The approach and methods presented here retain the simplicity of the PFT representation, but capture a wider range of functional diversity.

Materials and Methods

Data. The TRY database (www.try-db.org) (14) provided all data for leaf traits and the categorical traits to aggregate PFTs (TRY–Categorical Traits Dataset, https://www.try-db.org/TryWeb/Data.php#3, January 2016) used in the analysis. The TRY data may be requested from the TRY database custodians. See *SI Appendix*, section S10 for a complete list of the original publications associated with this subset of TRY. The extract from TRY used here has just under 45,000 measurements of individuals from 3,680 species with measurements of at least one of SLA, leaf nitrogen per dry leaf mass (N_m), and/or leaf phosphorus per leaf dry mass (P_m). The number of individual measurements varies from 32,315 for SLA on 2,953 species to 19,282 for N_m on 3,053 species down to 8,052 for P_m on 1,810 species; see *SI Appendix*, Table S4 for the number of unique measurements and species found in all categorizations used in the analysis. The species taxonomy was standardized using The Plant List (www.theplantlist.org/). Measurements were associated with environmental categories through Köppen–Geiger climate zones (44). All environmental variables are on a $0.5^\circ \times 0.5^\circ$ grid. Climate variables use 30-y climatologies from 1961 to 1990 as estimated by the Climate Research Unit (45, 46). Soil variables are from the International Soil Reference and Information Center–World Inventory of Soil Emission Potentials (ISRIC-WISE) (47). The spatial extent of PFTs has been previously estimated through satellite estimates of land cover around the year 2005 (48), and these estimates have been refined into climatic categories (15, 35). While TRY, and thus the data used here, represents the largest collection of plant traits in the world, most of the measurements come from a subset of global regions: North America, Europe, Australia, China, Japan, and Brazil. There are still large sections of the planet with extremely sparse measurements, notably much of the tropics outside of the Americas, large swaths of Central Asia, the Russian Federation, South Asia, and much of the Arctic (*SI Appendix*, Fig. S17). Improving data collection in these regions will greatly improve future modeling efforts. Until observations are more complete there remains the possibility of spurious patterns,

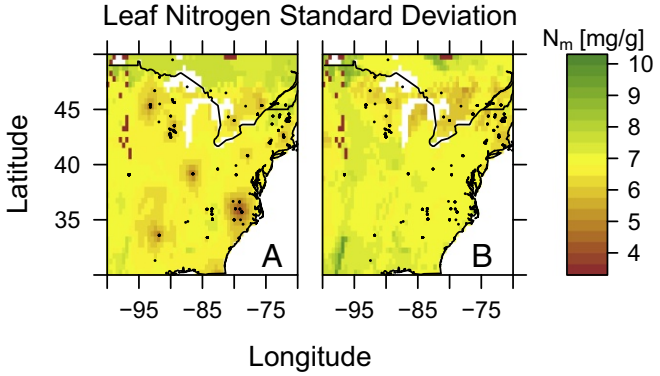

Fig. 6. Spatial learning. (*A*) The spatial model SD of N_m. The predicted variation near the data locations (black dots) is much lower than variation at locations away from any data point. (*B*) The linear model SD, which does not account for local spatial information, has no such pattern.

although we have found little evidence to suggest their presence in this analysis, even in comparison with detailed regional studies (*SI Appendix*, Fig. S26) (49).

Classification of PFTs and Categorical Model. We used three nested levels of PFT classification. In the first level, all plants are categorized into a single group ("PFT-free"). In the second level ("broad"), all plants are categorized into PFTs based on categorical traits associated with growth form (grass, shrub, tree) and leaf type (broad and needle-leaved), leading to the following four PFTs: grasses, shrubs, broad-leaved trees, and needle-leaved trees (*SI Appendix*, Fig. S1). In the third level ("narrow"), the broad PFTs are further refined by their climatic region—tropical, temperate, boreal—as well as leaf phenology and, for the grasses, photosynthetic pathway (C_3 or C_4). This produces 14 PFTs (*SI Appendix*, Fig. S2), which correspond exactly to those found in the CLM (4). Note that these PFT classifications exclude nonwoody eudicots ("herbs"), which were excluded from the analysis, on account of their lack of dominance within these PFT categories (50) and therefore, on account of being widely measured could overly influence the structure of the trait distributions if they were included. Satellite estimates of the PFT abundance that correspond to the narrow PFT categories defined above have already been calculated (15, 48) and we used these to assign a percentage of each $0.5^\circ \times 0.5^\circ$ pixel to each PFT present according to the fraction of the land surface within that pixel occupied by the PFT. The broad PFT fractions are calculated by summing the narrow PFT categories within each broad classification.

The categorical model uses the PFT categories and averages trait values for each species across individual measurements at each measured location. This defines the PFT as the interspecies range of trait values and ignores all local environmental factors. The results of the categorical model are summarized by the mean and SD of each PFT's trait values (*SI Appendix*, Table S4) for all three resolutions of the model. Note that in the PFT-free case where no PFT information is used, the categorical model produces a constant trait distribution across the entire vegetated world. The categorical model and the Bayesian models described in the following section all use location-specific species mean values to estimate trait distributions. We assume no intraspecific variation in trait values. However, in regions dominated by a small number of species this may lead to biased predictions. The hyperdominance of a small group of species in the Amazon has recently been demonstrated (51) and thus serves as a case study to evaluate our assumption of equal species weighting (*SI Appendix*, section S8, Fig. S23). We found that equal weights (species means) produced trait distribution estimates closest to those of the hyperdominant trait abundances and this reinforces the use of this assumption globally. Further, as noted above, the omission of herbaceous species from tropical regions in this analysis (and ref. 51) may unduly limit trait diversity and calls for further research.

Bayesian Models. A more fine-tuned depiction of geographical or spatial variation of plant trait values within each PFT can be achieved by leveraging environmental and location information, which allows trait values to adjust based on local conditions. Data for 17 climate- (45, 46) and soil-based (47) environmental predictors were available at the $0.5^\circ \times 0.5^\circ$-pixel resolution used to create the trait maps. To avoid overfitting and collinearity issues, these 17 predictors were screened (*SI Appendix*, section S7) based on correlations among predictors, based on their individual correlation with the traits, and to include climate covariates along different axes of environmental stress and both chemical and physical soil covariates. We finally selected 5 predictors—mean annual temperature (MAT), total annual radiation (RAD), moisture index (precipitation/evapotranspiration) (MI), percentage of hydrogen (aqueous) (pH), and percentage of clay content (CLY). Remote-sensing data products, such as Normalized Difference Vegetation Index (52), are not used as covariates, to allow for inference outside of the historical observation period through perturbations of environmental covariates.

We used environment–trait relationships to obtain predictions of trait values (1, 16–18, 37, 43) in a linear regression setup. The formal details of the initial model are as follows. We denote log-transformed trait values at a geographical location s as $y_{trait}(s)$. This set of five predictors at a location s is denoted by the vector $x(s) = (x_1(s), x_2(s), \ldots, x_5(s))'$. A linear regression model relating the trait to the environmental predictors is specified as

$$y_{trait}(s) = b_0 + b_1x_1(s) + b_2x_2(s) + \ldots + b_5x_5(s) + \epsilon(s), \quad [1]$$

where b_i are the regression coefficients and $\epsilon(s)$ is the error term explaining residual variation. Estimation of model parameters and prediction were achieved with a fully Bayesian hierarchical model. This enables inclusion of prior information and prediction of full trait distributions instead of representative values (like mean or median), thereby ensuring that the uncertainty associated with the estimation of model parameters is fully propagated into the predictive trait distributions.

We then generalized the above model into a Bayesian spatial linear regression model that borrows information from geographically proximal regions to capture residual spatial patterns beyond what is explained by environmental predictors. A customary specification of a spatial regression model is obtained by splitting up the error term $\epsilon(s)$ in Eq. **1** into the sum of a spatial process $w(s)$ and an error term $\eta(s)$ that accounts for the residual variation after adjusting for the spatial effects $w(s)$. The underlying latent process $w(s)$ accounts for local nuances beyond what is captured by the environmental predictors and is often interpreted as the net contribution from unobserved or unusable predictors. Gaussian processes (GPs) are widely used for modeling unknown spatial surfaces such as $w(s)$, due to their convenient formulation as a multivariate Gaussian prior for the spatial random effect, unparalleled predictive performance (53), and ease of generating uncertainty-quantified predictions at unobserved locations. We use the computationally effective nearest-neighbor GP (27), which nicely embeds into the Bayesian hierarchical setup as a prior for $w(s)$ in the second stage of the model specification. All technical specifications of the Bayesian spatial model are provided in *SI Appendix*, section S1.

The linear regression models used in previous studies (1, 16–18) and both the spatial and nonspatial Bayesian models described above assume a global relationship between the traits and environment. Given the goal of predicting trait values for the entire land surface, the assumption of a universal trait–environment relationship may be an oversimplification (54). Moreover, if there is significant variation in plant trait values among different PFTs, the estimated parameters will be skewed toward values from abundantly sampled PFTs, such as broad-leaved trees. Additional information about plant characteristics at a specific location, if available, can potentially be used to improve predictions. As mentioned earlier, we have PFT classifications for each observation of the dataset used here and satellite estimates of PFT abundance at all pixels. The global regression approaches described above ignore this information and can yield biased predictions at locations dominated by PFTs poorly represented in the data, such as shrubs. Hence, we also incorporate the PFT information in these regression models by allowing the trait–environment relationship to vary between different PFTs. Finally, the PFT-specific distributions from the Bayesian models were weighted by the satellite-based PFT abundances to create a landscape-scale trait distribution, thereby enabling straightforward comparison between all three categorizations of PFT. Details of the PFT-based Bayesian models are provided in *SI Appendix*, section S2. The use of a GP-based spatial model as well as the Bayesian implementation of the regression models was unique to this application of plant trait mapping and, as results indicated, were critical to improving model predictions as well as properly quantifying trait distributions.

ACKNOWLEDGMENTS. The authors appreciate the improvements suggested by two anonymous referees, which improved the clarity and depth of the manuscript. This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the US Department of Energy, Office of Science, Office of Biological and Environmental Research (Grant DE-SC0012677 to P.B.R. and A.B.). O.K.A. acknowledges the support of the Australian Research Council (CE140100008). This research was also funded by programs from the NSF Long-Term Ecological Research (Grant DEB-1234162) and Long-Term Research in Environmental Biology (Grant DEB-1242531). A.B., F.F., and P.B.R. acknowledge funding from NSF Grant IIS-1563950. P.B.R. also acknowledges support from two University of Minnesota Institute on the Environment discovery grants. This study has been supported by the TRY initiative on plant traits (www.try-db.org). The TRY database is hosted at the Max Planck Institute for Biogeochemistry (Jena, Germany) and supported by DIVERSITAS/Future Earth, the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, and the EU H2020 project BACI (Grant 640176). B.B. acknowledges a Natural Environment Research Council (NERC) independent research fellowship NE/M019160/1. J.P. acknowledges the financial support from the European Research Council Synergy Grant ERC-SyG-2013-610028 IMBALANCE-P, the Spanish Government Grant CGL2013-48074-P, and the Catalan Government Grant SGR 2014-274. B.B.-L. was supported by the Earth System Modeling program of the US Department of Energy, Office of Science, Office of Biological and Environmental Research. K.K. acknowledges the contribution of the Wageningen University and Research Investment theme Resilience for the project Resilient Forest (KB-29-009-003). P.M. acknowledges support from ARC Grant FT110100457 and NERC Grant NE/F002149/1. W.H. acknowledges support from the National Natural Science Foundation of China (Grant 41473068) and the "Light of West China" Program of the Chinese Academy of Sciences.

[a]Department of Forest Resources, University of Minnesota, St. Paul, MN 55108; [b]Department of Biostatistics, Johns Hopkins University, Baltimore, MD 21205; [c]Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN 55108; [d]Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455; [e]Australian Research Council Centre of Excellence in Plant Energy, Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia; [f]Division of Plant Sciences, Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia; [g]Max Planck Institute for Biogeochemistry, 07745 Jena, Germany; [h]German Centre for Integrative Biodiversity Research Halle-Jena-Leipzig, 04103 Leipzig, Germany; [i]UMR 1137 Ecologie et Ecophysiologie Forestières, Université de Lorraine–Institut National de la Recherche Agronomique, 54506 Vandoeuvre-lès-Nancy, France; [j]Environmental Change Institute, University of Oxford, Oxford OX1 3BJ, United Kingdom; [k]Joint Global Change Research Institute, Department of Energy Pacific Northwest National Laboratory, College Park, MD 20740; [l]Department of Geography and Geology, Kingston University London, Surrey KT1 2EE, United Kingdom; [m]School of Biological Sciences, Seoul National University, Seoul 08826, South Korea; [n]Plant Diversity and Ecosystems Management Unit, School of Biosciences & Veterinary Medicine, University of Camerino, 62032 Camerino, Italy; [o]Department of Theoretical and Applied Sciences, University of Insubria, I-21100 Varese, Italy; [p]Systems Ecology, Department of Ecological Science, Vrije Universiteit, 1081 HV Amsterdam, The Netherlands; [q]Jonah Ventures, Manhattan, KS 66502; [r]Department of Community Ecology, Helmholtz Centre for Environmental Research–UFZ, 06120 Halle (Saale), Germany; [s]School of Earth and Environmental Sciences, The University of Manchester, Manchester M13 9PT, United Kingdom; [t]Instituto Multidisciplinario de Biología Vegetal (Consejo Nacional de Invetigaciones Cientificas y Técnicas), Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba, CC 495 Córdoba, Argentina; [u]Departamento de Diversidad Biológica y Ecología, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba, CC 495 Córdoba, Argentina; [v]Faculdade de Filosofia Ciencias e Letras de Ribeirao Preto, Universidade de Sao Paulo, CEP 14040-901 Bairro Monte Alegre, Ribeirao Preto, Sao Paulo, Brazil; [w]Laboratory of Ecology Ecodiv, Institut National de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture, Normandie Université, 76821 Mont-Saint-Aignan, France; [x]Facultad de Ciencias Naturales y Matematicas, Universidad del Rosario, Bogota 110111, Colombia; [y]Departamento de Biología y Geología, Física y Química Inorgánica, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, 28933 Móstoles, Spain; [z]Institut National de la Recherche Agronomique, Unité Sous Contrat 1339, Centre d'Etude Biologique de Chizé, F 79360 Villiers en Bois, France; [aa]Centre D'Étude Biologique de Chizé, CNRS–Université La Rochelle (UMR 7372), F-79360 Villiers en Bois, France; [bb]College of Resources and Environmental Sciences, China Agricultural University, Beijing 100193, China; [cc]Key Laboratory of Biogeography and Bio-Resource in Arid Land, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, Xinjiang, China; [dd]School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, WITS 2050, Johannesburg, South Africa; [ee]Senckenberg Biodiversity and Climate Research Centre (BiK-F), 60325 Frankfurt/Main, Germany; [ff]Department of Physical Geography, Goethe-University, 60438 Frankfurt/Main, Germany; [gg]Institute of Systematic Botany and Ecology, Ulm University, 89081 Ulm, Germany; [hh]Team Vegetation, Forest and Landscape Ecology, Wageningen Environmental Research, 6708 PB Wageningen, The Netherlands; [ii]Chairgroup Forest Ecology and Management, Wageningen University, 6708 PB Wageningen, The Netherlands; [jj]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095; [kk]Department of Forest Vegetation, Forestry and Forest Products Research Institute, Tsukuba 305-8687, Japan; [ll]Department of Botany, University of Wyoming, Laramie, WY 82071; [mm]School of Geosciences, University of Edinburgh, Edinburgh EH9 3FF, United Kingdom; [nn]Institute of Biology and Environmental Science, University of Oldenburg, 26111 Oldenburg, Germany; [oo]Department of Plant Physiology, Estonian University of Life Sciences, 51014 Tartu, Estonia; [pp]Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan; [qq]CSIC, Unitat d'Ecologia Global CREAF-CSIC-UAB, Bellaterra 08193, Barcelona, Catalonia, Spain; [rr]CREAF, Cerdanyola del Vallès 08193, Barcelona, Catalonia, Spain; [ss]Department of Forestry, Michigan State University, East Lansing, MI 48824; [tt]Department of Biology, Algoma University, Sault Ste. Marie, ON P6A 2G4, Canada; [uu]Institute of Environmental Sciences, Leiden University, 2333 CC Leiden, The Netherlands; [vv]Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, CA 92521; [ww]Laboratório de Planejamento Ambiental, Embrapa Clima Temperado, Pelotas, RS, Brazil 96010-971; [xx]Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831; [yy]Museo Nacional de Ciencias Naturales, CSIC, E-28006 Madrid Spain; [zz]Department of Systematic Botany and Functional Biodiversity, University of Leipzig, 04103 Leipzig, Germany; and [aaa]Hawkesbury Institute for the Environment, Western Sydney University, Penrith NSW 2751, Australia

1. Van Bodegom PM, Douma JC, Verheijen LM (2014) A fully traits-based approach to modeling global vegetation distribution. *Proc Natl Acad Sci USA* 111:13733–13738.
2. Maire V, et al. (2015) Global effects of soil and climate on leaf photosynthetic traits and rates. *Glob Ecol Biogeogr* 24:706–717.
3. DeFries RS, et al. (1995) Mapping the land surface for global atmosphere-biosphere models: Toward continuous distributions of vegetation's functional properties. *J Geophys Res* 100:20867.
4. Bonan GB, et al. (2011) Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. *J Geophys Res* 116:1–22.
5. Reich PB, Ellsworth DS, Walters MB (1998) Leaf structure (specific leaf area) modulates photosynthesis–nitrogen relations : Evidence from within and across species and functional groups. *Funct Ecol* 12:948–958.
6. Kattge J, Knorr W, Raddatz T, Wirth C (2009) Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Glob Change Biol* 15:976–991.
7. Crous KY, et al. (2017) Nitrogen and phosphorus availabilities interact to modulate leaf trait scaling relationships across six plant functional types in a controlled-environment study. *New Phytol* 215:992–1008.
8. Wright IJ, et al. (2004) The worldwide leaf economics spectrum. *Nature* 428:821–827.
9. Reich PB, Oleksyn J, Wright IJ (2009) Leaf phosphorus influences the photosynthesis-nitrogen relation: A cross-biome analysis of 314 species. *Oecologia* 160: 207–212.
10. Atkin OK, et al. (2015) Global variability in leaf respiration in relation to climate, plant functional types and leaf traits. *New Phytol* 206:614–636.
11. Bahar N, et al. (2016) Leaf-level photosynthetic capacity in lowland Amazonian and high elevation, Andean tropical moist forests of Peru. *New Phytol* 214:1002–1018.
12. Rowland L, et al. (2016) Scaling leaf respiration with nitrogen and phosphorus in tropical forests across two continents. *New Phytol* 214:1064–1077.
13. Reich PB (2005) Global biogeography of plant chemistry: Filling in the blanks. *New Phytol* 168:263–266.
14. Kattge J, et al. (2011) TRY - a global database of plant traits. *Glob Change Biol* 17: 2905–2935.
15. Oleson KW, et al. (2013) Technical description of version 4.5 of the Community Land Model (CLM) (National Center for Atmospheric Research, Boulder, CO), Technical Report NCAR/TN-503+STR.
16. Reich PB, Oleksyn J (2004) Global patterns of plant leaf N and P in relation to temperature and latitude. *Proc Natl Acad Sci USA* 101:11001–11006.
17. Ordoñez JC, et al. (2009) A global study of relationships between leaf traits, climate and soil measures of nutrient fertility. *Glob Ecol Biogeogr* 18:137–149.
18. Simpson AH, Richardson SJ, Laughlin DC (2016) Soil-climate interactions explain variation in foliar, stem, root and reproductive traits across temperate forests. *Glob Ecol Biogeogr* 25:964–978.
19. Reich PB, Wright IJ, Lusk CH (2007) Predicting leaf physiology from simple plant and climate attributes: A global GLOPNET analysis. *Ecol Appl* 17:1982–1988.
20. Reich PB, Rich RL, Lu X, Wang YP, Oleksyn J (2014) Biogeographic variation in evergreen conifer needle longevity and impacts on boreal forest carbon cycle projections. *Proc Natl Acad Sci USA* 111:13703–13708.
21. Swenson NG, et al. (2012) The biogeography and filtering of woody plant functional diversity in North and South America. *Glob Ecol Biogeogr* 21:798–808.
22. Hawkins BA, Rueda M, Rangel TF, Field R, Diniz-Filho JAF (2014) Community phylogenetics at the biogeographical scale: Cold tolerance, niche conservatism and the structure of North American forests. *J Biogeogr* 41:23–38.
23. Šímová I, et al. (2015) Shifts in trait means and variances in North American tree assemblages: Species richness patterns are loosely related to the functional space. *Ecography* 38:649–658.
24. Swenson NG, et al. (2017) Phylogeny and the prediction of tree functional diversity across novel continental settings. *Glob Ecol Biogeogr* 26:553–562.
25. Douma JC, de Haan MWA, Aerts R, Witte JPM, van Bodegom PM (2012) Succession-induced trait shifts across a wide range of NW European ecosystems are driven by light and modulated by initial abiotic conditions. *J Ecol* 100:366–380.
26. Asner GP, Knapp DE, Anderson CB, Martin RE, Vaughn N (2016) Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proc Natl Acad Sci USA* 113:E4043–E4051.
27. Datta A, Banerjee S, Finley A, Gelfand A (2016) Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. *J Am Stat Assoc* 111:800–812.
28. Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. *Planta* 149:78–90.
29. Scheiter S, Higgins SI (2009) Impacts of climate change on the vegetation of Africa: An adaptive dynamic vegetation modelling approach. *Glob Change Biol* 15:2224–2246.
30. Scheiter S, Langan L, Higgins SI (2013) Next-generation dynamic global vegetation models: Learning from community ecology. *New Phytol* 198:957–969.
31. Pavlick R, Drewry DT, Bohn K, Reu B, Kleidon A (2012) The Jena Diversity-Dynamic Global Vegetation Model (JeDi-DGVM): A diverse approach to representing terrestrial biogeography and biogeochemistry based on plant functional trade-offs. *Biogeosciences Discussions* 9:4627–4726.

32. Pappas C, Fatichi S, Burlando P (2014) Terrestrial water and carbon fluxes across climatic gradients: Does plant diversity matter? *New Phytol* 16:3663.
33. Gross N, et al. (2017) Functional trait diversity maximizes ecosystem multifunctionality. *Nat Ecol Evol* 1:0132.
34. Clark DB, et al. (2011) The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics. *Geosci Model Dev* 4:701–722.
35. Bonan GB (2002) Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Glob Biogeochem Cycles* 16:5.1–5.18.
36. Meir P, Grace J, Miranda AC (2001) Leaf respiration in two tropical rainforests: Constraints on physiology by phosphorus, nitrogen and temperature. *Funct Ecol* 15:378–387.
37. Verheijen LM, et al. (2015) Inclusion of ecologically based trait variation in plant functional types reduces the projected land carbon sink in an earth system model. *Glob Chang Biol* 21:3074–3086.
38. Bloom AA, Exbrayat JF, van der Velde IR, Feng L, Williams M (2016) The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proc Natl Acad Sci USA* 113:1285–1290.
39. Meir P, Levy PE, Grace J, Jarvis PG (2007) Photosynthetic parameters from two contrasting woody vegetation types in West Africa. *Plant Ecol* 192:277–287.
40. Domingues TF, et al. (2010) Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. *Plant Cell Environ* 33:959–980.
41. Zhang Q, Wang YP, Pitman AJ, Dai YJ (2011) Limitations of nitrogen and phosphorous on the terrestrial carbon uptake in the 20th century. *Geophys Res Lett* 38:1–5.
42. Medlyn B, et al. (2016) Using models to guide field experiments: A priori predictions for the CO2 response of a nutrient- and water-limited native Eucalypt woodland. *Glob Change Biol* 22:2834–2851.
43. Verheijen LM, 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.
44. Peel B, Finlayson BL, McMahon TA (2007) Updated world map of the Köppen-Geiger climate classification. *Hydrol Earth Syst Sci* 11:1633–1644.
45. New M, Hulme M, Jones P (1999) Representing twentieth-century space–time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *J Clim* 12:829–856.
46. Harris I, Jones PD, Osborn TJ, Lister DH (2014) Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *Int J Climatol* 34:623–642.
47. Batjes NH (2005) ISRIC-WISE global data set of derived soil properties on a 0.5 by 0.5 degree grid (Version 3.0) (World Soil Information, Wageningen, The Netherlands), Report 2005/08.
48. Lawrence PJ, Chase TN (2007) Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0). *J Geophys Res* 112:G01023.
49. Asner GP et al. (2017) Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science* 355:385–389.
50. Gibson DJ (2009) *Grasses & Grassland Ecology* (Oxford Univ Press, New York).
51. ter Steege H, et al. (2013) Hyperdominance in the Amazonian tree flora. *Science* 342:1243092.
52. Ollinger SV, et al. (2008) Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proc Natl Acad Sci USA* 105:19336–19341.
53. Rasmussen C (1996) *Evaluation of Gaussian processes and other methods for non-linear regression*. PhD thesis (University of Toronto, Toronto).
54. Verheijen LM, Aerts R, Bönisch G, Kattge J, Van Bodegom PM (2016) Variation in trait trade-offs allows differentiation among predefined plant functional types: Implications for predictive ecology. *New Phytol* 209:563–575.

Supplementary Materials: Mapping local and global variability in plant trait distributions

S1 Bayesian Spatial Model Details

In this section we present the details of the Bayesian spatial model described in the main text. For illustrative purposes we use SLA to define all the models, the models for N_m and P_m are identical. Let $y(\mathbf{s})$ denote the log-transformed value of SLA response observed at a location $\mathbf{s}$. Additionally, let $\mathbf{x}(\mathbf{s})$ denote the 5×1 vector of the environmental predictors used in the regression model at the location $\mathbf{s}$ i.e. $\mathbf{x}(\mathbf{s}) = (\mathrm{MAT}(\mathbf{s}), \mathrm{RAD}(\mathbf{s}), \mathrm{MI}(\mathbf{s}), \mathrm{pH}(\mathbf{s}), \mathrm{CLY}(\mathbf{s}))'$. The Bayesian spatial linear regression model relating the response to the covariates at the location $\mathbf{s}$ is given by

$$\mathbf{y}(\mathbf{s}) = \alpha + \mathbf{x}(\mathbf{s})'\boldsymbol{\beta} + w(\mathbf{s}) + \epsilon(\mathbf{s}) \tag{1}$$

In the above equation, α denotes the intercept, $\boldsymbol{\beta} = (\beta_{MAT}, \beta_{MI}, \beta_{RAD}, \beta_{pH}, \beta_{CLY})'$ denotes the regression coefficients, $w(\mathbf{s})$ denotes the spatial random effect to be specified later and $\epsilon(\mathbf{s})$ denotes independent and identically distributed noise which explains the residual variation after accounting for the predictors and the spatial effect. Note that we used species specific averages of raw trait values at each location as observations to eliminate species level sampling bias. So, if a trait measurement is recorded for r species at a given location, we have r observations for that location, one for each species. Let $\mathbf{y}$ denote the $n \times 1$ vector stacking up all the observations of SLA.

Similarly we denote the full design matrix as $\mathbf{X}$, the full spatial vector as $\mathbf{w}$ and the noise vector as $\boldsymbol{\epsilon}$. Since, some locations have multiple observations, these sets of observations will share the same spatial random effect $\mathbf{w}(\mathbf{s})$. We denote this smaller set of n_0 unique locations by $\mathcal{S}$ and the corresponding vector of spatial random effects as $\mathbf{w}_{\mathcal{S}}$. Thus we have $\mathbf{w} = \mathbf{M}\mathbf{w}_{\mathcal{S}}$ where $\mathbf{M}$ is a $n \times n_0$ binary matrix mapping each observation to its unique spatial location. Gaussian Processes are widely used as priors for spatial random effects. We use the recently proposed Nearest Neighbor Gaussian Process (NNGP) prior (Datta et al. 2016) due its computational efficiency i.e. we assume $\mathbf{w}(\mathbf{s}) \sim NNGP_m(\mathbf{0}, \widetilde{C}(\cdot,\cdot \,|\, \sigma^2, \phi, \nu))$ where $\widetilde{C}$ denotes the NNGP covariance function derived using m nearest neighbors from a Matérn covariance function C (Cressie and Wikle 2011) with marginal variance σ^2, decay parameter ϕ and smoothness ν. Following the data analysis in Datta et al. (2016), we used $m = 5$ neighbors to construct the NNGP and chose $\nu = 0.5$ corresponding to the exponential covariance function. This implies that the spatial effect vector $\mathbf{w}_{\mathcal{S}}$ follows a Normal distribution with zero mean and NNGP covariance matrix $\widetilde{\mathbf{C}}_{\mathcal{S}}$. We also assume normality of the residual errors i.e. $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \tau^2\mathbf{I})$. The full Bayesian model specification is completed by adding priors for the regression and covariance parameters. We assume conjugate Normal priors for the regression coefficient and Inverse-Gamma (IG) priors for the variances σ^2 and τ^2. The range parameter ϕ was assigned an uniform ($Unif$) prior. The full model specification is thus given by:

$$N(\mathbf{y} \,|\, \mathbf{X}\boldsymbol{\beta} + \mathbf{M}\mathbf{w}_{\mathcal{S}}, \tau^2\mathbf{I}) \times N(\mathbf{w}_{\mathcal{S}} \,|\, \mathbf{0}, \widetilde{\mathbf{C}}_{\mathcal{S}}) \times N(\boldsymbol{\beta} \,|\, \boldsymbol{\mu}_\beta, \mathbf{V}_\beta) \times \\ IG(\sigma^2 \,|\, a_\sigma, b_\sigma) \times IG(\tau^2 \,|\, a_\tau, b_\tau) \times Unif(\phi \,|\, a_\phi, b_\phi). \tag{2}$$

We used a flat prior for $\boldsymbol{\beta}$ with $\boldsymbol{\mu}_\beta = \mathbf{0}$ and $\mathbf{V}_\beta^{-1} = 10^{-5}\mathbf{I}$. We chose the Inverse Gamma shape parameters $a_\sigma = a_\tau = 2$ and the rate parameters $b_\sigma = 0.2$ and $b_\tau = 0.1$. This implies that a priori we have $E(\sigma^2) = 0.2$, $E(\tau^2) = 0.1$ and $Var(\sigma^2) = Var(\tau^2) = \infty$. The lower and upper limits for the Uniform prior on ϕ were chosen to be 50 and 2000. This implies that the effective range (radius within which spatial correlation is greater than 0.05) varies widely from 10 to 384 kilometers. Posterior distributions of all the parameters and spatial random effects were obtain using a Markov

Chain Monte Carlo sampler outlined in Datta et al. (2016).

S2 PFT based model

We incorporate the available information about Plant Functional Types into the Bayesian spatial model through two different models.

Broad PFT based model: Each observation is classified as belonging to one of the four broad PFT groups – Needleleaf (N), Broadleaf (B), Shrubs (S) and Grasses (G). Subsequently we have 4 vectors of observations $\mathbf{y}_N$, $\mathbf{y}_B$, $\mathbf{y}_S$ and $\mathbf{y}_G$ with dimensions n_N, n_B, n_S and n_G respectively where $n_N + n_B + n_S + n_G = n$. In this Broad PFT based approach, we assume that the regression coefficients differ for each PFT. Consequently we denote the design matrix for the group x as $\mathbf{X}_x$ and the corresponding regression coefficient vector as $\boldsymbol{\beta}_x$ where $x = N, B, S, G$. We also assume PFT specific spatial random effects for the model. Let $\mathcal{S}(x)$ denote the set of unique locations where some data was observed for the broad PFT x and $\mathbf{w}^{(x)}(\mathbf{s})$ denote the corresponding spatial process where $x = N, B, S, G$. Finally, if $\mathbf{M}_x$ denotes the binary matrix mapping the location of i^{th} observation in $\mathbf{y}_x$ to $\mathcal{S}(x)$, then the regression model is specified as:

$$\mathbf{y}_x = N\left(\mathbf{X}_x\boldsymbol{\beta}_x + \mathbf{M}_x\mathbf{w}^{(x)}_{\mathcal{S}(x)}, \tau_x^2\mathbf{I}\right) \text{ for } x \in \{N, B, S, G\} \tag{3}$$

The priors for the spatial random effects, covariance and regression parameters are similar to the specifications in Section S1 except that each parameter is replaced by a set of four parameters one corresponding to each broad PFT.

Narrow PFT based model: Within each broad PFT, each observation can be further classified into narrow PFTs. For the broad PFT classification, we estimated the spatial regression model parameters, described in Section S1, within each of the four PFT categories. Adoption of a similar approach for the narrow PFT classification was thwarted by a lack of data within each of the fourteen PFTs. For example, the TRY dataset has observations in 922 distinct pixels for Pm.

However, some of the fourteen narrow PFTs had fewer than 10 pixels. Such low sample sizes are inadequate to estimate the six regression coefficients (including the intercept) let alone the spatial process parameters. Instead, the model incorporates a PFT specific intercept for each narrow PFT to allow for variation in the overall trait level between these 14 PFTs. The slope parameters representing the trait-environment relationships were assumed to be the same for all narrow PFTs within each of the four broad PFTs, circumventing the low sample size. Let broad PFT group x contain narrow PFTs $1, 2, \ldots, k$. Further, let $I_k(\mathbf{s})$ denote the binary indicator if an observation $y^{(x)}(\mathbf{s})$ from broad PFT group x at location $\mathbf{s}$ belongs to the narrow PFT k. Then the narrow PFT based model is given by:

$$y^{(x)}(\mathbf{s}) = \sum_{i=1}^{k} \alpha_i^{(x)} I_k(\mathbf{s}) + \mathbf{x}^{(x)}(\mathbf{s})'\beta^{(x)} + w^{(x)}(\mathbf{s}) + \epsilon^{(x)}(\mathbf{s}) \text{ for } x \in \{N, B, S, G\} \tag{4}$$

where $\alpha_i^{(x)}$ denotes the PFT specific intercept. All parameters are assigned priors similar to the PFT-free and broad PFT models.

Table S1: Parameters estimates (posterior median and 95% confidence intervals in braces) from the narrow spatial model for SLA. The means and standard deviations for the five environmental predictors used to derive the respective Z-scores are as follows: MAT (8.4 ± 14), RAD (1616 ± 525), MI (0.95 ± 0.7), pH (6.3 ± 0.9) and CLY (23.5± 6.7).

	Needleleafs	Broadleafs	Shrubs	Grasses
PFT1	1.609 (1.468, 1.763)	NA	NA	NA
PFT2	1.719 (1.571, 1.874)	NA	NA	NA
PFT3	2.08 (1.689, 2.467)	NA	NA	NA
PFT4	NA	2.4 (2.197, 2.608)	NA	NA
PFT5	NA	2.242 (2.142, 2.336)	NA	NA
PFT6	NA	2.514 (2.305, 2.727)	NA	NA
PFT7	NA	2.684 (2.586, 2.772)	NA	NA
PFT8	NA	2.797 (2.711, 2.885)	NA	NA
PFT9	NA	NA	2.138 (2.001, 2.274)	NA
PFT10	NA	NA	2.634 (2.488, 2.775)	NA
PFT11	NA	NA	2.719 (2.589, 2.858)	NA
PFT12	NA	NA	NA	2.778 (2.312, 3.277)
PFT13	NA	NA	NA	2.833 (2.708, 2.929)
PFT14	NA	NA	NA	2.901 (2.761, 3.017)
MAT	0.12 (-0.216, 0.451)	-0.056 (-0.209, 0.107)	0.037 (-0.245, 0.32)	0.085 (-0.16, 0.333)
RAD	0.052 (-0.163, 0.285)	-0.111 (-0.227, -0.01)	-0.182 (-0.373, 0.008)	-0.191 (-0.393, 0.01)
MI	0.053 (-0.024, 0.132)	0.018 (-0.019, 0.053)	0.029 (-0.027, 0.086)	0.001 (-0.069, 0.069)
pH	-0.049 (-0.172, 0.086)	-0.046 (-0.106, 0.019)	-0.155 (-0.25, -0.062)	-0.069 (-0.187, 0.049)
CLY	0.081 (-0.039, 0.206)	0.042 (-0.014, 0.106)	0.104 (-0.007, 0.217)	0.055 (-0.067, 0.178)
τ^2	0.053 (-0.024, 0.132)	0.018 (-0.019, 0.053)	0.029 (-0.027, 0.086)	0.001 (-0.069, 0.069)
σ^2	0.118 (0.072, 0.185)	0.088 (0.07, 0.112)	0.099 (0.065, 0.145)	0.136 (0.095, 0.195)
ϕ	58 (50, 117)	131 (52, 948)	978 (623, 1959)	61 (51, 116)
range	332 (164, 382)	146 (20, 370)	20 (10, 312)	313 (166, 380)

Table S2: Parameters estimates (posterior median and 95% confidence intervals in braces) from the narrow spatial model for N_m. The means and standard deviations for the five environmental predictors used to derive the respective Z-scores are as follows: MAT (8.4 ± 14), RAD (1616 ± 525), MI (0.95 ± 0.7), pH (6.3 ± 0.9) and CLY (23.5± 6.7).

	Needleleafs	Broadleafs	Shrubs	Grasses
PFT1	2.456 (2.399, 2.513)	NA	NA	NA
PFT2	2.532 (2.484, 2.58)	NA	NA	NA
PFT3	2.569 (2.433, 2.704)	NA	NA	NA
PFT4	NA	2.823 (2.701, 2.95)	NA	NA
PFT5	NA	2.726 (2.653, 2.794)	NA	NA
PFT6	NA	2.931 (2.805, 3.057)	NA	NA
PFT7	NA	2.938 (2.873, 3.003)	NA	NA
PFT8	NA	3.092 (3.026, 3.158)	NA	NA
PFT9	NA	NA	2.63 (2.512, 2.751)	NA
PFT10	NA	NA	2.917 (2.79, 3.046)	NA
PFT11	NA	NA	3.12 (3.008, 3.226)	NA
PFT12	NA	NA	NA	2.965 (2.584, 3.37)
PFT13	NA	NA	NA	2.848 (2.776, 2.918)
PFT14	NA	NA	NA	2.677 (2.591, 2.76)
MAT	0.138 (0.031, 0.25)	0.118 (0.017, 0.226)	0.205 (-0.007, 0.415)	0.099 (-0.035, 0.225)
RAD	-0.066 (-0.142, 0.01)	-0.124 (-0.193, -0.05)	-0.114 (-0.266, 0.033)	-0.13 (-0.252, 0.005)
MI	-0.021 (-0.05, 0.006)	-0.002 (-0.032, 0.027)	-0.012 (-0.089, 0.065)	-0.03 (-0.084, 0.021)
pH	-0.006 (-0.049, 0.034)	0.051 (0.004, 0.089)	0.046 (-0.049, 0.136)	0.044 (-0.035, 0.123)
CLY	0 (-0.04, 0.04)	0.005 (-0.033, 0.049)	0.032 (-0.051, 0.118)	-0.007 (-0.075, 0.062)
τ^2	0.032 (0.028, 0.038)	0.068 (0.065, 0.071)	0.081 (0.068, 0.097)	0.081 (0.073, 0.09)
σ^2	0.021 (0.015, 0.028)	0.041 (0.034, 0.051)	0.06 (0.041, 0.089)	0.073 (0.054, 0.099)
ϕ	53 (50, 67)	52 (50, 63)	65 (50, 1325)	55 (50, 83)
range	365 (286, 383)	367 (303, 383)	296 (14, 381)	349 (232, 383)

Table S3: Parameters estimates (posterior median and 95% confidence intervals in braces) from the narrow spatial model for P_m. The means and standard deviations for the five environmental predictors used to derive the respective Z-scores are as follows: MAT (8.4 ± 14), RAD (1616 ± 525), MI (0.95 ± 0.7), pH (6.3 ± 0.9) and CLY (23.5± 6.7).

	Needleleafs	Broadleafs	Shrubs	Grasses
PFT1	0.127 (0.014, 0.235)	NA	NA	NA
PFT2	0.195 (0.113, 0.282)	NA	NA	NA
PFT3	0.155 (-0.064, 0.381)	NA	NA	NA
PFT4	NA	0.263 (0.034, 0.511)	NA	NA
PFT5	NA	0.202 (0.061, 0.338)	NA	NA
PFT6	NA	0.405 (0.173, 0.655)	NA	NA
PFT7	NA	0.467 (0.334, 0.599)	NA	NA
PFT8	NA	0.557 (0.435, 0.681)	NA	NA
PFT9	NA	NA	-0.174 (-0.476, 0.11)	NA
PFT10	NA	NA	0.284 (-0.024, 0.581)	NA
PFT11	NA	NA	0.661 (0.47, 0.851)	NA
PFT12	NA	NA	NA	0.265 (-0.351, 0.841)
PFT13	NA	NA	NA	0.373 (0.247, 0.49)
PFT14	NA	NA	NA	0.469 (0.318, 0.613)
MAT	-0.115 (-0.302, 0.075)	-0.224 (-0.425, -0.017)	-0.087 (-0.512, 0.331)	0.019 (-0.207, 0.248)
RAD	-0.098 (-0.24, 0.053)	0.019 (-0.121, 0.16)	0.09 (-0.229, 0.427)	-0.208 (-0.437, 0.027)
MI	-0.056 (-0.108, -0.004)	-0.083 (-0.143, -0.025)	0.108 (-0.123, 0.334)	-0.105 (-0.23, 0.015)
pH	-0.079 (-0.155, -0.002)	0.002 (-0.084, 0.089)	0.033 (-0.178, 0.23)	-0.052 (-0.211, 0.11)
CLY	0.005 (-0.069, 0.082)	-0.028 (-0.104, 0.052)	0.035 (-0.15, 0.219)	0.056 (-0.067, 0.182)
τ^2	0.057 (0.046, 0.072)	0.115 (0.109, 0.123)	0.181 (0.136, 0.245)	0.174 (0.15, 0.204)
σ^2	0.071 (0.048, 0.099)	0.121 (0.094, 0.154)	0.149 (0.082, 0.245)	0.145 (0.093, 0.227)
ϕ	55 (50, 81)	61 (50, 114)	939 (64, 1940)	822 (79, 1955)
range	348 (237, 382)	316 (169, 381)	20 (10, 299)	23 (10, 244)

S3 Model Summary Statistics

The typical summary statistics for the log-normal distribution are the location, μ and shape σ parameters which correpond to the mean, m, and standard deviation, s calculated on a log-transformed normal distribution. However, the un-transformed parameters are closely related:

$$\mu = \ln\left(\frac{m}{\sqrt{1+\frac{s^2}{m^2}}}\right), \tag{5}$$

and similarly

$$\sigma = \sqrt{\ln\left(1+\frac{s^2}{m^2}\right)}. \tag{6}$$

So, for consistency with prior work we present m and s as a full description of the distribution. Furthermore we note in the limit of a true log-normal distribution μ is the median of the un-transformed values. Neither set of summary statistics are as satisfying for the log-normal as the mean and standard deviation are for a standard normal distribution, which is why we recommend analysis on the full distribution. Below, we report the means and standard deviations for the categorical models reported in the manuscript.

Table S4: Mean, Standard Deviation, and number of species within each classification: PFT-free, broad PFTs, and narrow PFTs. Abbreviations indicate each of the narrow PFTs which are preceded by their respective broad PFTs as indicated by square brackets.
[Needleleaf trees] NETte: Needleleaf Evergreen Tree, temperate; NETb: Needleleaf Evergreen Tree, boreal; NDTb: Needleleaf Deciduous Tree, boreal;
[Broadleaf trees] BETtr: Broadleaf Evergreen Tree, tropical; BETte: Broadleaf Evergreen Tree, temperate; BDTtr: Broadleaf Deciduous Tree, tropical; BDTte: Broadleaf Deciduous Tree, temperate; BDTb: Broadleaf Deciduous Tree, boreal;
[Shrubs] BESte: Broadleaf Evergreen Shrub, temperate; BDSte: Broadleaf Deciduous Shrub, temperate; BDSb: Broadleaf Deciduous Shrub, boreal;
[Grasses] C3ar: C3 arctic grass; C3: C3 grass; C4: C4 grass.

PFT Category (# individuals/species)	SLA [$m^2 kg^{-1}$]	N_m [mg N g^{-1}]	P_m [mg P g^{-1}]
PFT-free (44645/3680)	14.5±7.5	20.1±7.1	1.3±0.8
Needleleaf Trees (4758/124)	6.3±3.5	12.5±3.4	1.2±0.6
Broadleaf Trees (29582/3061)	14.2±6.9	20.7±6.9	1.2±0.7
Shrubs (3315/312)	13.8±8	20.5±8.6	1.7±1.1
Grasses (6990/503)	18.6±9.4	18±6.8	1.8±0.9
NETte (2051/62)	5.7±2.6	11.8±3.5	1.2±0.8
NETb (2625/56)	6.7±4.2	13±3.1	1.3±0.4
NDTb (82/6)	9.5±0.3	14.6±4.4	1.4±0.5
BETtr (15759/1739)	13.9±6.2	20.2±6.7	1±0.5
BETte (2609/548)	10.8±5.9	18.5±7.2	1.3±0.8
BDTtr (2413/288)	16.2±6.3	23.5±7	1.2±0.6
BDTte (4540/319)	17±8.5	21.8±7	1.8±1
BDTb (4261/167)	18.5±8.2	24.1±5.3	2±0.7
BES (1076/154)	10±5.5	18.2±9.2	1.1±0.7
BDSte (1353/95)	16.5±8.6	22.4±7	1.8±1.2
BDSb (886/63)	17.4±8.2	22.5±8.4	2.2±1
C3ar (78/19)	17.5±7.4	19.4±6.4	1.4±0.7
C3 (5283/274)	19.2±9.9	19.9±7.1	1.8±0.9
C4 (1629/210)	17.5±8.6	15.6±5.8	1.7±0.9

S4 Evaluation Metrics

The coefficient of determination (R^2) is a widely used regression statistic to determine the goodness of fit in linear models. The R^2 statistic provides, on a scale of 0 to 1, a quantitative estimate of the improvement achieved by introduction of predictors in a classical regression setup relative to the baseline (constant mean) model. However, this classical definition of the R^2 is meaningful only for evaluating linear least squares models. In the Bayesian paradigm, an equivalent statistic does not exist. Popular Bayesian model selection methods include Deviance Information Criterion (DIC – (Spiegelhalter et al. 2002)) as well as the scoring rules proposed by Gelfand and Ghosh (1998) and Gneiting and Raftery (2007). These metrics are typically not confined to any particular range and are hence difficult to interpret in absolute terms. Moreover, they are only applicable for Bayesian models, whereas we sought a common platform to evaluate the fits from the categorical model and the posterior means from the Bayesian model. To this end, we resorted to an alternative interpretation of R^2, where under certain model assumptions, it is equivalent to the squared correlation between the observed values and the model fitted values. This definition can be used to generate an R^2 statistic for any model output. It is important to note that posterior estimates from the Bayesian models typically do not conform to the assumptions required for the equivalent definition to hold. Nonetheless, the new definition is very interpretable on its own and values close to one still imply a good fit to the training data. Moreover, it provides the desired common platform for comparing the frequentist approach for the categorical model and the Bayesian spatial model. We refer to this variant of R^2 as pseudo-R^2 to differentiate it from its classical analogue.

The R^2 statistic has been widely criticized for its predilection for more inclusive and complex linear models. Spiess and Neumeyer (2010) demonstrated, via extensive simulation experiments, how R^2 (or even adjusted R^2) performs very poorly in this regard. It is easy to illustrate, both theoretically and empirically, that linear models with large number of predictors tend to yield high R^2 despite little or no association with the response of interest. Hence, a close fit to the training data does not necessarily correlate with good predictive performance. The pseudo-R^2 defined above

also inherits the same pitfalls. Since our main goal is to create predictive maps of plant traits, we evaluated the predictive performance of all the models using Root Mean Square Predictive Error, RMSPE (Yeniay and Goktas 2002). The RMSPE calculation proceeds by holding out a randomly selected part of the data and using the predictions from the estimated model for the holdout dataset to assess the predictive performance. Formally it is defined as:

$$\text{RMSPE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_{hi} - \widehat{y_{hi}})^2} \tag{7}$$

where $(y_{h1}, y_{h2}, \ldots, y_{hn})$ denotes the holdout data, and $(\widehat{y_{h1}}, \widehat{y_{h2}}, \ldots, \widehat{y_{hn}})$ are the corresponding fits from the model estimated from the training data. RMSPE is at the same scale as the trait value and can thus be interpreted as the expected prediction error when predicting at a new location. Unlike R^2, RMSPE does not necessarily favor more complex models. Models with lower RMSPE tend to have better predictive performance.

The Bayesian approach generates the posterior distributions of plant traits at every location. This transition from mean estimation to distribution estimation seamlessly delivers inference about plant trait variance, skewness, tail behavior and any other interesting aspect of the full trait distribution. The models ability to accurately predict trait distributions was evaluated with the Bayesian predictive credible intervals. We use the same hold out locations used for computing RMSPE to calculate the predictive Coverage Probability (CP). CP is defined as the percentage of holdout locations covered by the respective 95% Bayesian credible intervals. Formally, for every hold out observation y_{hi}, let $\widehat{y_{hi,low}}$ and $\widehat{y_{hi,up}}$ denote the lower and upper 2.5% quantiles from the posterior predictive sample for y_{hi}. Then, the interval $(\widehat{y_{hi,low}} - \widehat{y_{hi,up}})$ constitutes a 95% posterior credible interval for y_{hi}. If the posterior sample provides a good estimate of the trait density at this location, then y_{hi} is expected to lie within a 95% posterior confidence interval with 95% probability. CP is

defined as the coverage probability averaged over all the holdout data i.e.,

$$\text{CP} = \frac{1}{n}\sum_{i=1}^{n} I(y_{hi} \in (\widehat{y_{hi,low}}, \widehat{y_{hi,up}})) \tag{8}$$

where I denotes the binary function indicating the inclusion or exclusion of y_{hi} in its credible interval. A CP value close to 95% indicates good coverage. CP values close to one indicate too wide credible intervals resulting from overestimation of the tail while low CP values indicate overfitting to the training data and thereby underestimating the trait variance.

S5 Additional data analysis results

Here are the goodness of fit metrics from 20 different holdout sets, but only 100 samples in each pixel. The average model metrics and their respective standard deviations (parentheses) are presented in Table S5. The overall pattern is similar to Table 1 of the main article. The Kolmogorov-Smirnov statistics from the case studies are then summarized in Table S6.

Table S5: Each model is indicated by a two letter abbreviation: C=Categorical, L=Linear, S=Spatial and the accompanying PFT resolution: f=PFT-free, b=broad, n=narrow.

Model	pseudo-R^2	RMSPE	CP
SLA			
Cf	NA (NA)	8.21 (0.697)	0.93 (0.014)
Cb	0.19 (0.009)	7.62 (0.664)	0.95 (0.013)
Cn	0.28 (0.012)	7.3 (0.623)	0.94 (0.012)
Lf	0.05 (0.006)	8.12 (0.626)	0.92 (0.018)
Lb	0.25 (0.007)	7.51 (0.56)	0.92 (0.015)
Ln	0.33 (0.01)	7.27 (0.578)	0.93 (0.015)
Sf	0.46 (0.008)	7.79 (0.703)	0.94 (0.015)
Sb	0.6 (0.008)	7.13 (0.636)	0.94 (0.009)
Sn	0.61 (0.009)	6.99 (0.653)	0.94 (0.012)
N_m			
Cf	NA (NA)	7.3 (0.404)	0.95 (0.013)
Cb	0.12 (0.005)	6.95 (0.438)	0.94 (0.013)
Cn	0.2 (0.006)	6.73 (0.396)	0.93 (0.011)
Lf	0.05 (0.01)	7.27 (0.439)	0.93 (0.015)
Lb	0.17 (0.008)	6.99 (0.415)	0.92 (0.014)
Ln	0.24 (0.008)	6.77 (0.411)	0.92 (0.014)
Sf	0.44 (0.006)	6.88 (0.493)	0.94 (0.013)
Sb	0.54 (0.006)	6.52 (0.407)	0.94 (0.011)
Sn	0.55 (0.006)	6.43 (0.445)	0.93 (0.013)
L_m			
Cf	NA (NA)	0.86 (0.081)	0.93 (0.023)
Cb	0.05 (0.009)	0.85 (0.081)	0.92 (0.021)
Cn	0.28 (0.012)	0.77 (0.077)	0.92 (0.026)
Lf	0.24 (0.013)	0.81 (0.082)	0.92 (0.025)
Lb	0.29 (0.015)	0.79 (0.079)	0.91 (0.029)
Ln	0.34 (0.01)	0.78 (0.079)	0.91 (0.026)
Sf	0.63 (0.011)	0.79 (0.084)	0.92 (0.02)
Sb	0.68 (0.01)	0.79 (0.075)	0.93 (0.021)
Sn	0.69 (0.009)	0.79 (0.085)	0.93 (0.021)

Table S6: Kolmogorov-Smirnov (K-S) Statistic for Barro Colorado Island (BCI) and 18 aggregated pixels of the United States Southwest (USSW) for all nine models and three traits. The lowest value and best match is highlighted in bold. Each model is indicated by a two letter abbreviation: C=Categorical, L=Linear, S=Spatial and the accompanying PFT resolution: f=PFT-free, b=broad, n=narrow.

Model	SLA [BCI]	N_m [BCI]	P_m [BCI]	SLA [USSW]	N_m [USSW]	P_m [USSW]
Cf	0.211	0.244	0.281	0.255	0.291	0.348
Cb	0.214	0.207	0.224	0.241	0.266	0.327
Cn	0.205	0.205	0.243	0.261	0.267	**0.097**
Lf	0.337	0.261	0.297	0.240	0.135	0.172
Lb	0.298	0.278	0.210	0.155	**0.120**	0.114
Ln	0.237	0.215	0.210	0.189	0.175	0.108
Sf	0.163	0.119	**0.160**	0.173	0.150	0.109
Sb	**0.118**	0.134	0.190	**0.160**	0.179	0.117
Sn	**0.118**	**0.093**	0.220	0.184	0.220	0.156

S6 Training and Test data

Trait data from 85% of the locations were used as training data, while the remaining randomly selected 15% were withheld as test data for evaluation. The set of locations withheld were different for the three traits owing to their differential geographical coverage. For each trait, the same set of withheld locations were used to evaluate both the categorical and the Bayesian spatial models. We used stratified sampling to obtain the withheld locations. Stratification was based on three zones: tropics (all points between the Tropics of Capricorn and Cancer), temperate (extending from the Tropic of Cancer to the Arctic Circle and from the Tropic of Capricorn to the Antarctic Circle) and boreal (north of the Arctic Circle, there are no data south of the Antarctic Circle). The stratified sampling strategy ensured that each of the geographical zones were equi-proportionally represented in the withheld locations. Consequently, most of the 14 narrow PFTs were also significantly represented in the withheld sample.

S7 Covariate Selection

The 17 environmental predictors at each of the 53,900 pixels contain redundant information as indicated in Figure S18, which plots the pairwise correlations among the predictors and clearly indicates collinearity among them. In particular, the four temperature variables: yearly mean monthly maximum temperature (mxT), yearly mean monthly minimum temperature (miT), mean annual (yearly) temperature (MAT or myT) and mean temperature of the warmest quarter ($mTwq$) were highly correlated. Similarly, there was strong correlation among the precipitation indices: precipitation (pre), number of wet days (wet) and moisture index (MI). Among the soil covariates, total nitrogen ($totn$) and the carbon/nitrogen ratio (cnr) demonstrated strong correlation. Similar strong associations were detected between organic carbon ($orgC$) and bulk density (blk) and again among silt (slt), sand (snd) and clay (cly) content. Overall there was substantial collinearity among the covariates which may result in unreliable regression estimates. Surrogating the individual

predictors with the first few score vectors obtained from a principal component analysis is one solution to alleviate such collinearity (Dormann et al. 2013). However, such an approach was less attractive, in this case, as it is difficult to understand the trait-environment relationship from the principal component regression.

It is also important to note that the range of values of the environmental predictors in the training dataset is only a small subset of the observed combinations of values. This poor predictor coverage is an artifact of the low data density in many regions of the world. To elucidate the lack of coverage, we plotted the relative ranges of all the predictors in Figure S19. Note that for many predictors, the observed range only covers about 70% of the possible range. Consequently, a global regression model will estimate the trait-environment relationships from a much more limited set than the full span of global values into which it will predict. This problem is exacerbated by a large number of predictors, as the joint relative coverage (ratio of the range of values of the predictors in the observed locations to the range of values in the entire predicted space) of different environmental combinations is exponentially reduced with increased dimensionality of the predictor space. This is illustrated in Figures S20 and S21, where the pairwise joint relative coverages for the climate and soil predictors are significantly worse than the individual relative coverages. Using a large number of predictors in the training dataset will yield over-fitted models, resulting in a potentially inaccurate estimation of the trait-environment relationship at locations with environmental combinations not well represented in the data.

Together, collinearity and poor coverage of the predictors warrants a reduction in model complexity to ensure reliable predictions at new locations. General model selection methods like the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) are computationally infeasible in this setting as there are 2^{17} different predictor combinations to choose from. Over the last two decades, the statistical community has developed a large inventory of model selection techniques for analyzing complex datasets with ever-growing predictor dimensionality. Popular methods include the Lasso (Tibshirani 1994), SCAD (Fan and Li 2001), Dantzig Selector (Candès

and Tao 2007), and Adaptive Lasso (Zou 2006) methods, among others. Bayesian alternatives include stochastic search variable selection (George and McCulloch 1993), Bayesian Lasso (Park and Casella 2008), horseshoe prior (Carvalho et al. 2010) and others. Some approaches are also tailored to accommodate grouped selection of collinear predictors, such as the elastic net (Zou and Hastie 2005). However, model selection in a Bayesian spatial linear regression model is relatively less developed. Furthermore, none of the aforementioned approaches are designed to incorporate the available predictor data at unobserved locations. To the best of our knowledge, there is no extant work on predictor selection approaches in a Bayesian spatial paradigm that uses the information about the coverage of the predictors in the observed and unobserved data to infer the optimal choice of predictors.

Shorthanded by the dearth of a relevant model selection technique, we adopted a heuristic variable screening approach predicated upon pairwise trait-predictor correlation, predictor coverage and coverage over different axes of environmental sensitivity. For consistent interpretation of the model output we also restricted ourselves to the same subset of predictors for all three traits. Figure S22 shows the absolute values of the pairwise trait-predictor correlations between the seventeen predictors and three traits. We observed that among the four temperature variables miT (mean annual minimum temperature) and myT (mean annual temperature) exhibited relatively high correlations with all three traits. In addition, from Figure S19, we observed that all the temperature variables possessed poor relative coverage. Since the four temperature covariates were highly correlated among each other (S18), only one temperature-based predictor was retained in the model. We preferred myT over miT because it has been more widely used in other regression models (Reich and Oleksyn 2004; Ordoñez et al. 2009) and facilitates comparison of this work with others. The number of days above 0° C - *dg0* - was not included because it was highly correlated with myT, which was already included in the model and *dg0* has the lowest absolute correlation with N_m. Among the three correlated precipitation variables, *wet*, *pre* and *MI*, we observed that *wet* had the worst coverage ratio among all seventeen predictors (S19). It also has the lowest correlation

with SLA (S22,top). We chose *MI* (yearly precipitation/Penman-Monteith potential evapotranspiration) over *pre* because although the latter had slightly higher pairwise correlation with all the traits, *MI* had the best coverage ratio (97%) among all seventeen predictors (*pre* had 70% coverage ratio). Furthermore, *MI* better represents the moisture availability of a region by accounting for both inputs (precipitation) and demands (potential evapotranspiration). Radiation, *rad*, was also included in the model because it had a high correlation with all three traits and rounds out the three physical climate axes: temperature, moisture, and radiation. Among the soil predictors, *pH*— was almost uncorrelated with any of the other predictors. It also had high correlation with the traits and decent coverage (80%), and is a good representation of the general nutrient availability in the soil (Binkley et al. 1989; Maire et al. 2015). The next four soil variables *totn*, *orgC*, *cnr* and *blk* were not included due to strong collinearity, poor association with the traits and poor coverage. The last three soil variables all represent physical properties of the soil *slt*, *snd* and *cly* were all collinear. Among them, *slt* and *cly* both had strong correlation with the traits. We chose the latter because it had slightly better correlations with the traits and significantly better coverage (82% compared to 66% for slt).

To summarize, the initial suite of seventeen candidate predictors was narrowed down to five predictor variables based on the strength of their individual correlations and the representative coverage within TRY locations of the total global distribution of the predictor. This means that inference into areas without predictor values in the training data was minimized. These predictors were then standardized into z-scores by subtracting their mean values and dividing by their standard deviations. The five selected environmental predictors include three climate and two soil variables. Climate variables were *myT* (mean yearly temperature), *MI* (moisture index), and *rad* (annual total radiation). The soil variables selected were *pH*, as an aggregate of nutrient availability, and *cly* (percent clay content), for a physical soil characteristic.

S8 Amazonian Hyperdominance

Recently, ter Steege et al. (2013) reported that a small subset of species accounts for a large portion of tropical trees by stem count. Failing to account for this in the data analysis can introduce potential bias into the trait distributions at regions where such hyperdominance of species is prevalent. Ideally, if information on species abundance were available globally and at a fine resolution, those abundances could be used to weight the trait data for our analysis, and hence account for the species level bias in the data. In the absence of such a global database, we have addressed this potential bias issue at a coarser resolution - at the PFT level. This was done by reweighting the PFT specific trait distributions at each location with the remote sensing estimates of PFT abundance corresponding to that location. While this mitigates PFT level bias it does not address species level bias within a PFT.

Without knowledge of the true species abundance within each PFT at a location, we give equal weights to all species measured at that location. The other option would have been to use the TRY species abundance at that location as surrogate weights. This relies on the assumption that the species abundance in the TRY database mirrors the true abundance worldwide, which is hard to verify. However, the recent work on hyperdominance in tropical communities gives us an opportunity to evaluate our choice of weights against this alternative.

We tested whether weighting our data by the TRY occurence of the 227 most common species in the Amazon improved our trait distribution estimates. The trait distribution obtain by using the true abundance weights of the 227 species as reported in ter Steege et al. (2013) is used as the gold standard against which the two weighting schemes are compared. Figure S23 plots the three distributions each of SLA, N_m and P_m for the broadleaf tropical PFT across the Amazon, the region of study in (ter Steege et al. 2013). The black curve corresponds to trait distribution based on the 227 hyper abundant species with the abundance weights chosen to be those reported in ter Steege et al. (2013). The dotted lines in the figure correspond to the respective modes of each distribution.

We found that the results which most closely matched the trait distributions weighted by the

species occurrence in ter Steege et al. (2013) were the unweighted distributions (red curve). Despite the large proportion of TRY trait data coming from the hyperdominant species (20 to 30% of all measurements), the TRY weights for all species did not improve the trait distribution (blue curve) estimates. The modes of the trait distributions corroborates this as the mode from the equal-weighted scheme aligned much better with the gold standard mode. This suggests that given the available data the unweighted species level trait estimates are the best indicator of trait diversity.

Both choices of weights produced trait distributions whose broad shapes were very similar to the respective true abundance weighted distributions. This may reflect that within the narrow PFTs, the bias in the distribution of species may not necessarily translate to a bias of equivalent magnitude in the distribution of traits. Until we have both improved global level species location data accompanied by substantial intra-species trait measurements the assumption of equal weighting across species is, apparently, the best predictor available.

References

G. P. Asner, R. E. Martin, D. E. Knapp, R. Tupayachi, C. B. Anderson, F. Sinca, N. R. Vaughn, and W. Llactayo. Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science*, 355(6323):385–389, 2017. ISSN 0036-8075. doi: 10.1126/science.aaj1987. URL `http://www.sciencemag.org/lookup/doi/10.1126/science.aaj1987`.

Dan Binkley, C. Driscoll, H.L. Allen, P. Schoeneberge, and D. McAvoy. *Acidic Deposition and Forest Soils: Context and Case Studies in the Southeastern U.S.* Springer-Verlag Ecological Studies #72, New York, 1989.

E. Candès and T. Tao. The dantzig selector: Statistical estimation when p is much larger than n. *Annals of Statistics*, 35:2313–2351, 2007.

Carlos M. Carvalho, Nicholas G. Polson, and James G. Scott. The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480, 2010. doi: 10.1093/biomet/asq017.

Noel A. C. Cressie and Christopher K. Wikle. *Statistics for spatio-temporal data*. Wiley Series in Probability and Statistics. Wiley, Hoboken, NJ, 2011. ISBN 978-0-471-69274-4. URL `http://opac.inria.fr/record=b1133266`.

Abhirup Datta, Sudipto Banerjee, Andrew O. Finley, and Alan E. Gelfand. Hierarchical nearest-neighbor gaussian process models for large geostatistical datasets. *Journal of the American Statistical Association*, 111(514):800–812, 2016. doi: 10.1080/01621459.2015.1044091. URL `http://dx.doi.org/10.1080/01621459.2015.1044091`.

Carsten F. Dormann, Jane Elith, Sven Bacher, Carsten Buchmann, Gudrun Carl, Gabriel Carré, Jaime R García Marquéz, Bernd Gruber, Bruno Lafourcade, Pedro J. Leitão, Tamara Münkemüller, Colin Mcclean, Patrick E. Osborne, Björn Reineking, Boris Schröder, Andrew K. Skidmore, Damaris Zurell, and Sven Lautenbach. Collinearity: A review of methods to deal with

it and a simulation study evaluating their performance. *Ecography*, 36(1):027–046, 2013. ISSN 09067590. doi: 10.1111/j.1600-0587.2012.07348.x.

J. Fan and R. Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96:1348—1360, 2001.

A. E. Gelfand and S. K. Ghosh. Model choice: A minimum posterior predictive loss approach. *Biometrika*, 85:1–11, 1998.

Edward I. George and Robert E. McCulloch. Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88(423):881–889, 1993. doi: 10.1080/01621459.1993.10476353. URL `http://www.tandfonline.com/doi/abs/10.1080/01621459.1993.10476353`.

T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378, 2007.

Vincent Maire, Ian J. Wright, I. Colin Prentice, Niels H. Batjes, Radika Bhaskar, Peter M. van Bodegom, Will K. Cornwell, David Ellsworth, Ülo Niinemets, Alejandro Ordonez, Peter B. Reich, and Louis S. Santiago. Global effects of soil and climate on leaf photosynthetic traits and rates. *Global Ecology and Biogeography*, 24(6):706–717, 2015. ISSN 14668238. doi: 10.1111/geb.12296.

Jenny C. Ordoñez, Peter M. Van Bodegom, Jan Philip M Witte, Ian J. Wright, Peter B. Reich, and Rien Aerts. A global study of relationships between leaf traits, climate and soil measures of nutrient fertility. *Global Ecology and Biogeography*, 18(2):137–149, 2009. ISSN 1466822X. doi: 10.1111/j.1466-8238.2008.00441.x.

Trevor Park and George Casella. The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008. doi: 10.1198/016214508000000337. URL `http://dx.doi.org/10.1198/016214508000000337`.

P B Reich and J Oleksyn. Global patterns of plant leaf N and P in relation to temperature and latitude. *Proceedings of National Academy of Sciences*, 101(30):11001–11006, 2004.

David J. Spiegelhalter, Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, 64: 583–639, 2002.

A-N Spiess and N. Neumeyer. An evaluation of r^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a monte carlo approach. *BMC pharmacology*, 10: 6, 2010.

H. ter Steege, N.C.A. Pitman, D. Sabatier, C. Baraloto, R.P. Salomão, J.E. Guevara, O.L. Phillips, C.V. Castilho, W.E. Magnusson, J.F. Molino, A. Monteagudo, P. Núñez Vargas, J.C. Montero, T.R. Feldpausch, E.N.H. Coronado, T.J. Killeen, B. Mostacedo, R. Vasquez, R.L. Assis, J. Terborgh, F. Wittmann, A. Andrade, W.F. Laurance, S.G.W. Laurance, B.S. Marimon, B.H. Marimon, I.C. Guimarães Vieira, I.L. Amaral, R. Brienen, H. Castellanos, D. Cárdenas López, J.F. Duivenvoorden, H.F. Mogollón, F.D.A. Matos, N. Dávila, R. García-Villacorta, P.R. Stevenson Diaz, F. Costa, T. Emilio, C. Levis, J. Schietti, P. Souza, A. Alonso, F. Dallmeier, A.J.D. Montoya, M.T. Fernandez Piedade, A. Araujo-Murakami, L. Arroyo, R. Gribel, P.V.A. Fine, C.A. Peres, M. Toledo, G.A. Aymard C, T.R. Baker, C. Cerón, J. Engel, T.W. Henkel, P. Maas, P. Petronelli, J. Stropp, C.E. Zartman, D. Daly, D. Neill, M. Silveira, M.R. Paredes, J. Chave, D.A. Lima Filho, P.M. Jørgensen, A. Fuentes, J. Schöngart, F. Cornejo Valverde, A. Di Fiore, E.M. Jimenez, M.C. Peñuela Mora, J.F. Phillips, G. Rivas, T.R. van Andel, P. von Hildebrand, B. Hoffman, E.L. Zent, Y. Malhi, A. Prieto, A. Rudas, A.R. Ruschel, N. Silva, V.A. Vos, S. Zent, A.A. Oliveira, A.C. Schutz, T. Gonzales, M. Trindade Nascimento, H. Ramirez-Angulo, R. Sierra, M. Tirado, M.N. Umaña Medina, G. van der Heijden, C.I.A. Vela, E. Vilanova Torre, C. Vriesendorp, O. Wang, K.R. Young, C. Baider, H. Balslev, C. Ferreira, I. Mesones, A. Torres-Lezama, L.E. Urrego Giraldo, R. Zagt, M.N. Alexiades, L. Hernandez, I. Huamantupa-Chuquimaco, W. Milliken, W. Palacios Cuenca, D. Pauletto, E. Valderrama Sandoval, L. Valenzuela Gamarra, K.G. Dexter, K. Feeley, G. Lopez-Gonzalez, and M.R. Silman. Hyperdomi-

nance in the Amazonian tree flora. *Science*, 342(6156):1243092, 2013. ISSN 1095-9203. doi: 10.1126/science.1243092.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–288, 1994.

O. Yeniay and A. Goktas. A comparison of partial least squares regression with other prediction methods. *Hacettepe Journal of Mathematics and Statistics*, 31:99–111, 2002.

H. Zou. The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101:1418–1429, 2006.

H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67:301–320, 2005.

S9 Supplemental Figures

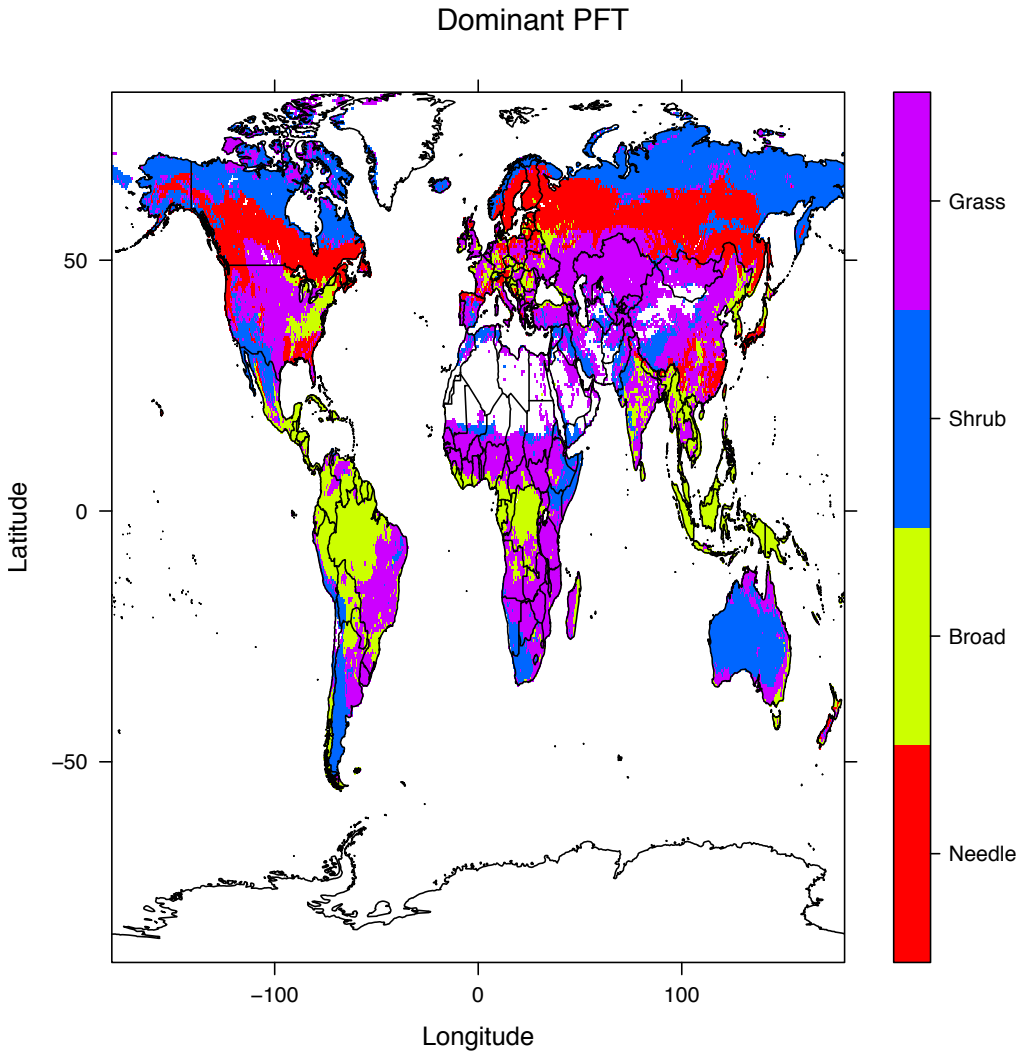

Figure S1: Dominant PFTs in Broad PFT categories. Note that multiple Broad PFTs are typically present, but for ease of reading only the PFT that covers the largest area in each pixel is displayed on the map.

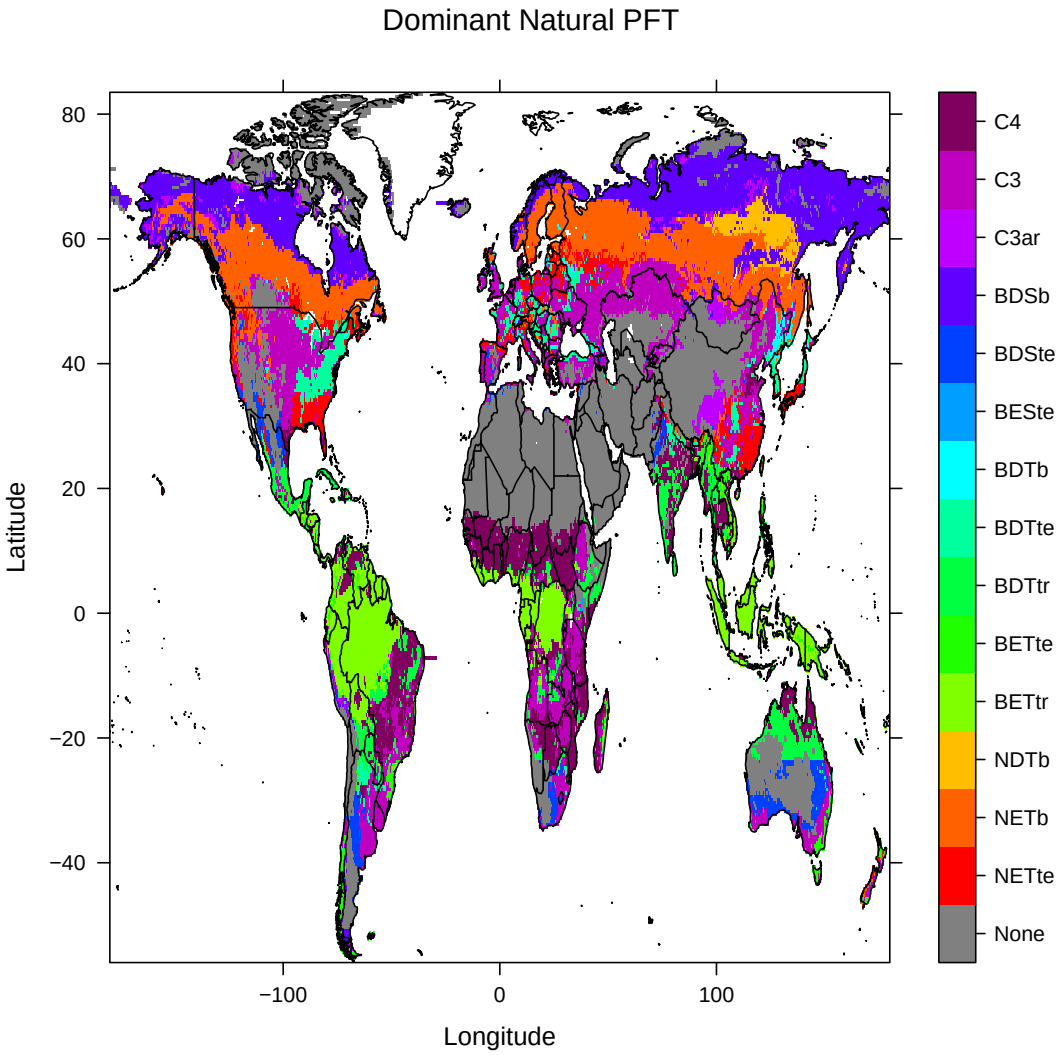

Figure S2: Dominant natural PFT in each pixel. Each pixel contains a mixture of PFTs, but for simplicity only the PFT occupying the largest fraction of the pixel is shown here. None indicates regions dominated by bare soil, rock, and/or crops. Abbreviations indicate each of the narrow PFTs which are preceded by their respective broad PFTs as indicated by square brackets.
[Needleleaf trees] NETte: Needleleaf Evergreen Tree, temperate; NETb: Needleleaf Evergreen Tree, boreal; NDTb: Needleleaf Deciduous Tree, boreal;
[Broadleaf trees] BETtr: Broadleaf Evergreen Tree, tropical; BETte: Broadleaf Evergreen Tree, temperate; BDTtr: Broadleaf Deciduous Tree, tropical; BDTte: Broadleaf Deciduous Tree, temperate; BDTb: Broadleaf Deciduous Tree, boreal;
[Shrubs] BESte: Broadleaf Evergreen Shrub, temperate; BDSte: Broadleaf Deciduous Shrub, temperate; BDSb: Broadleaf Deciduous Shrub, boreal;
[Grasses] C3ar: C3 arctic grass; C3: C3 grass; C4: C4 grass.

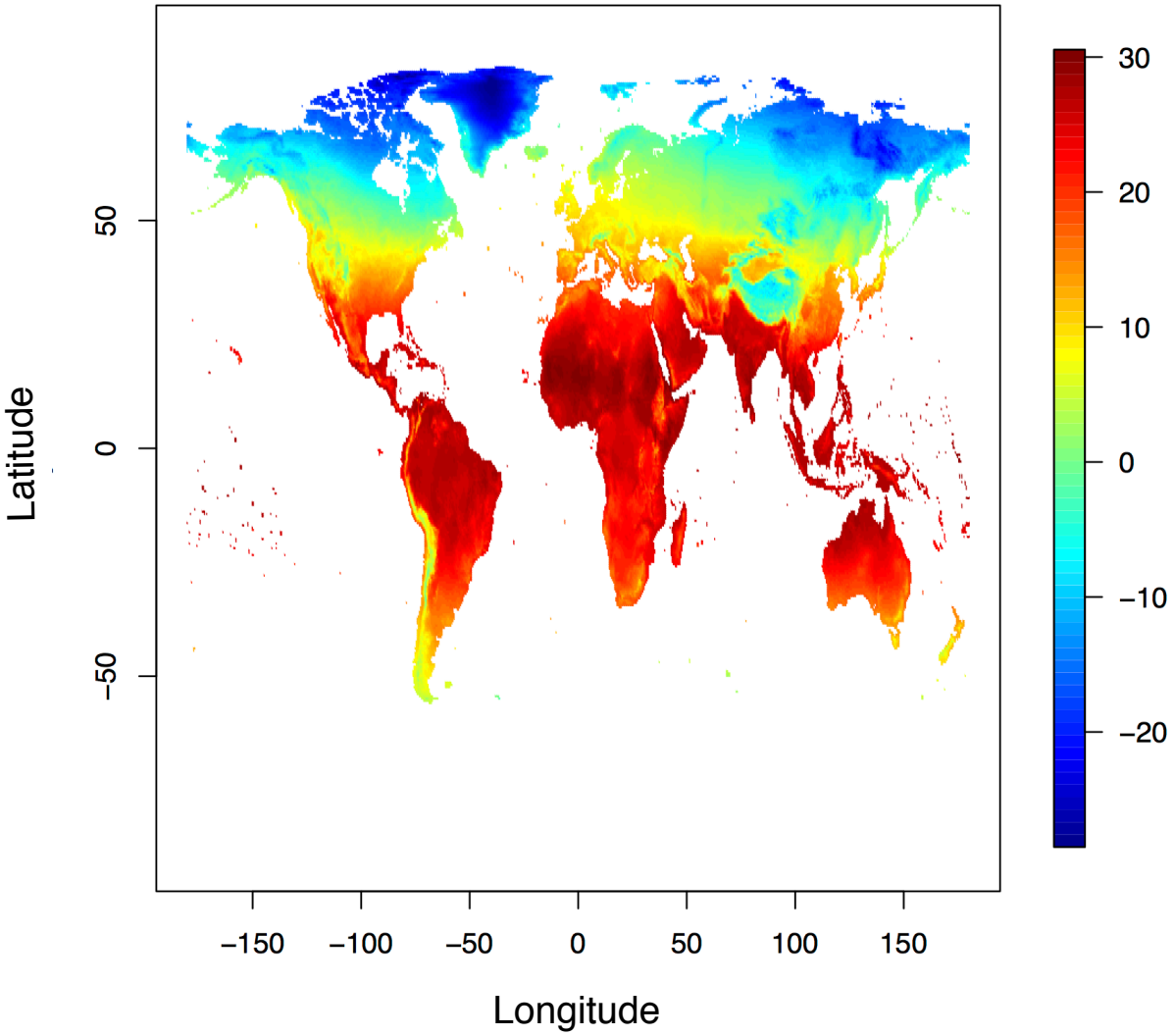

Figure S3: Mean Annual Temperature in Degrees Celsius.

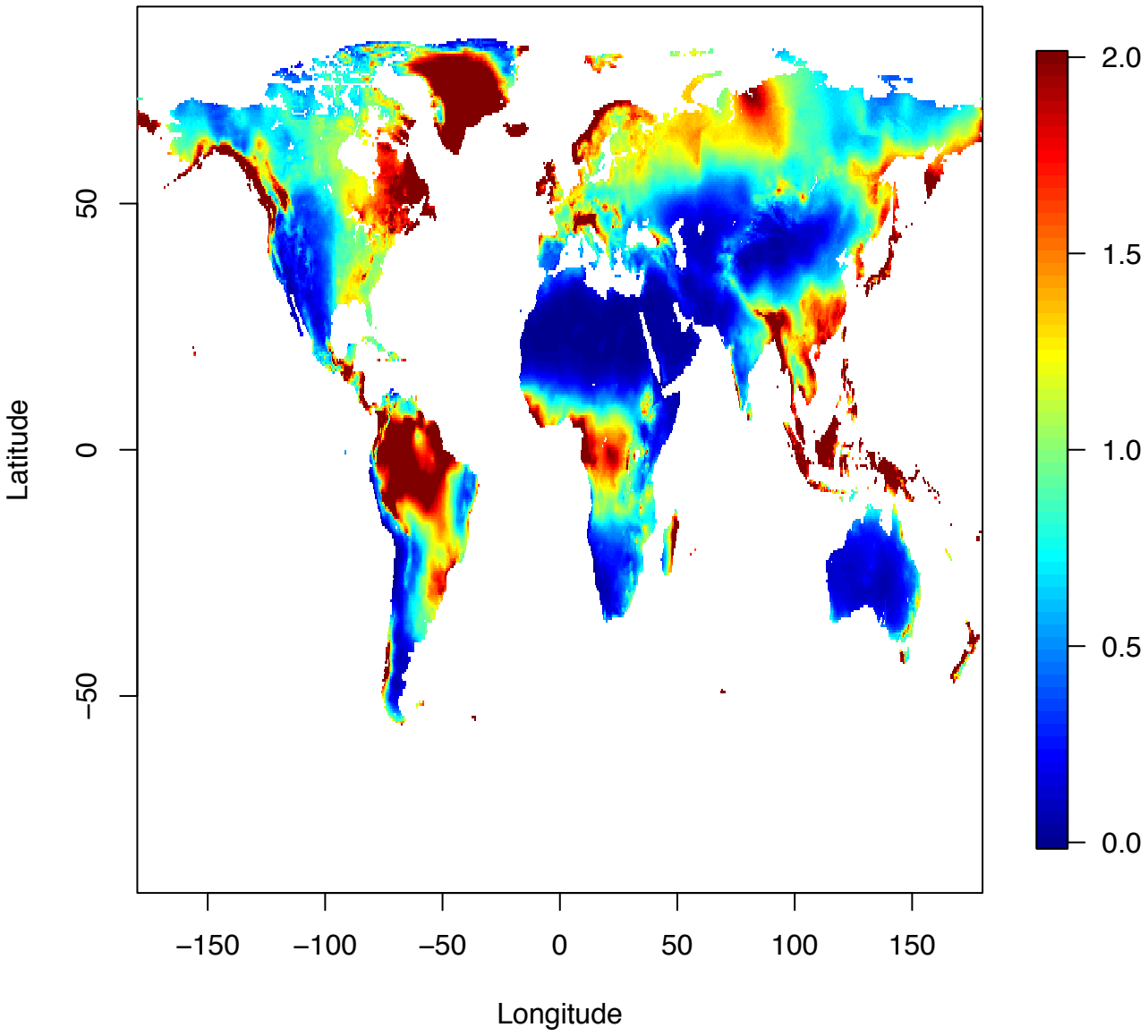

Figure S4: Moisture Index (Yearly Precipitation/Yearly Penman-Monteith Potential Evapotranspiration)

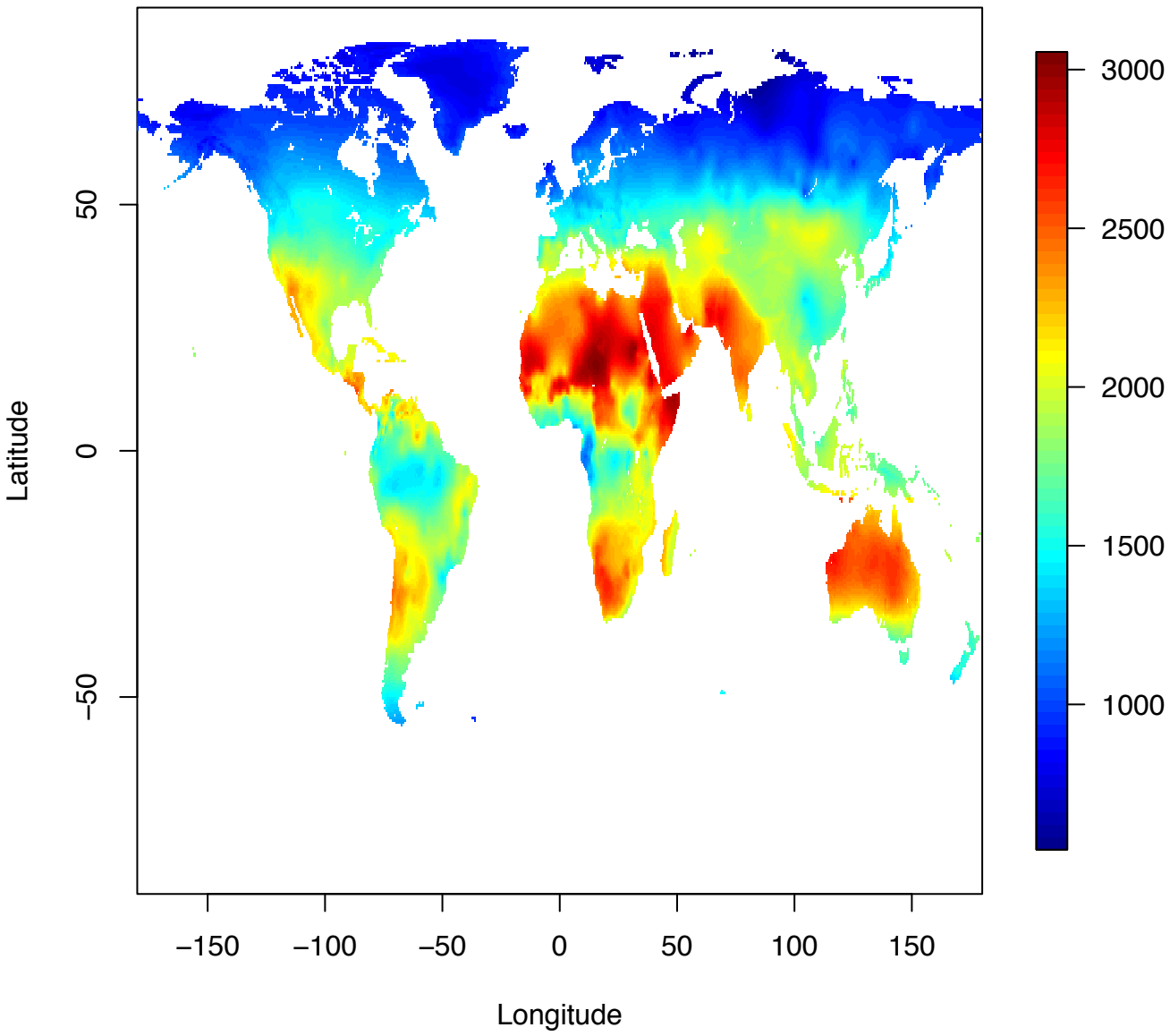

Figure S5: Yearly total radiation in units of W m^{-2}

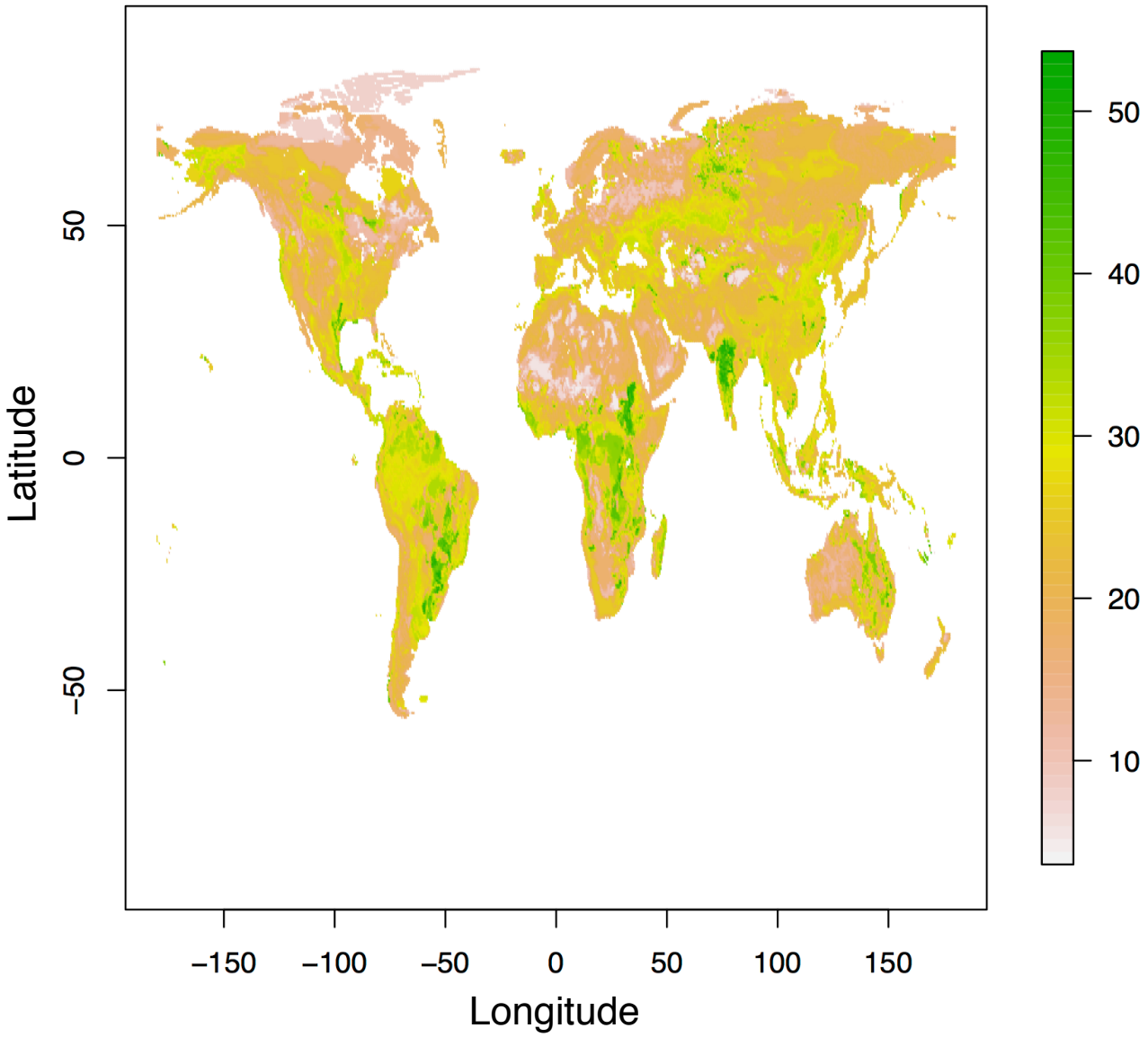

Figure S6: Soil clay content, in percent

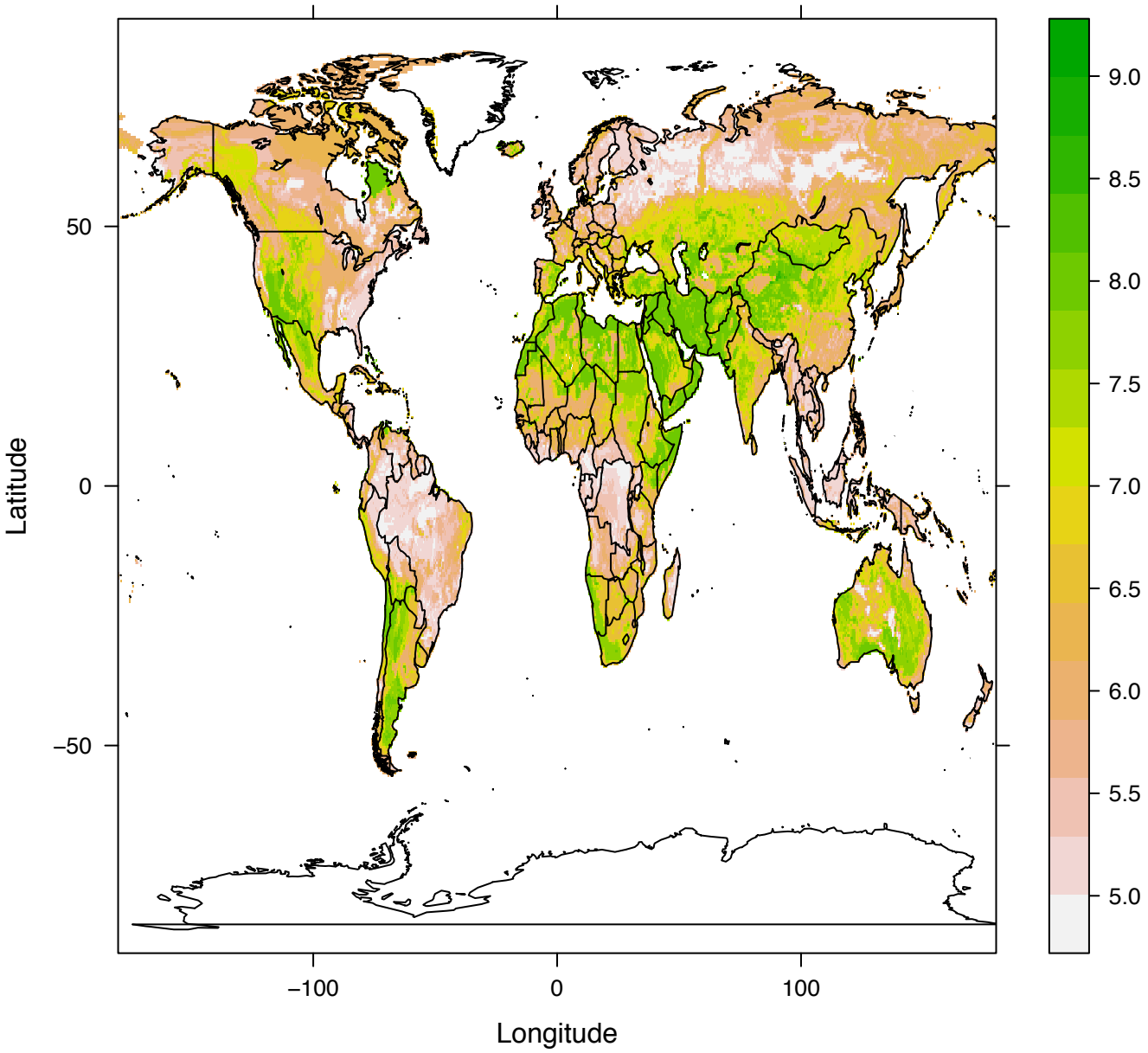

Figure S7: Soil pH

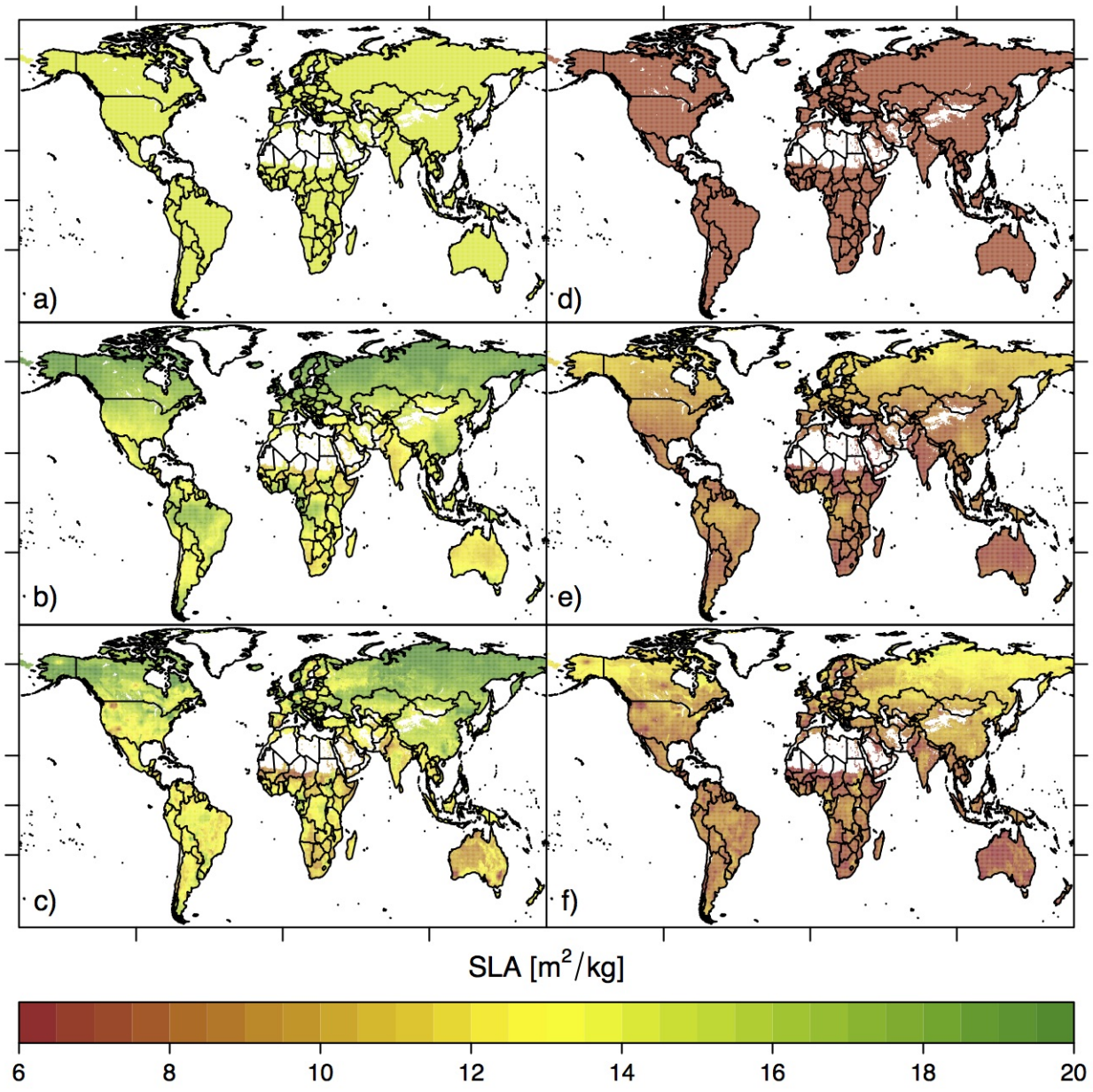

Figure S8: PFT-free SLA Models [m^2 kg^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

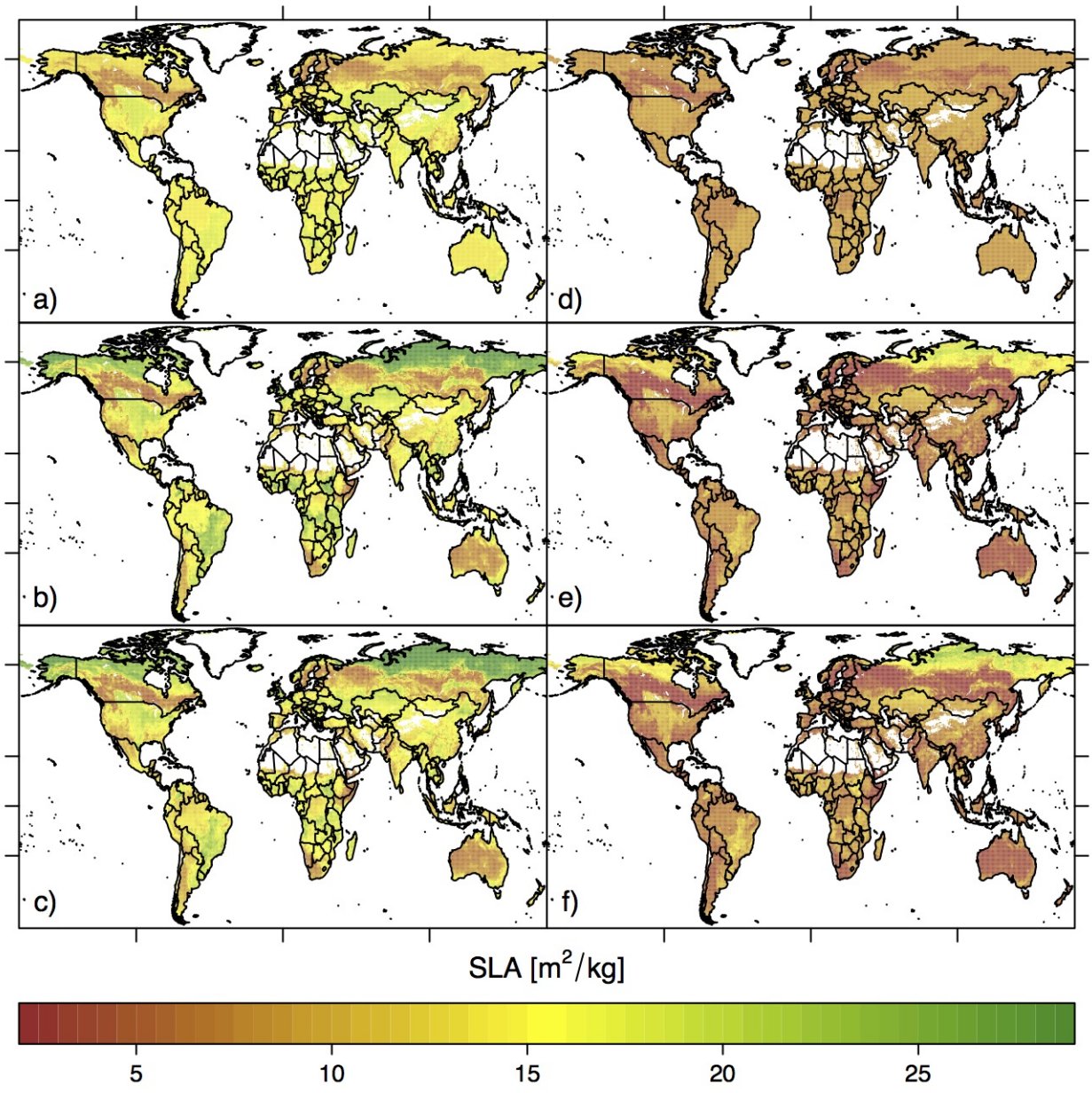

Figure S9: Broad PFT SLA Models [m^2 kg^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

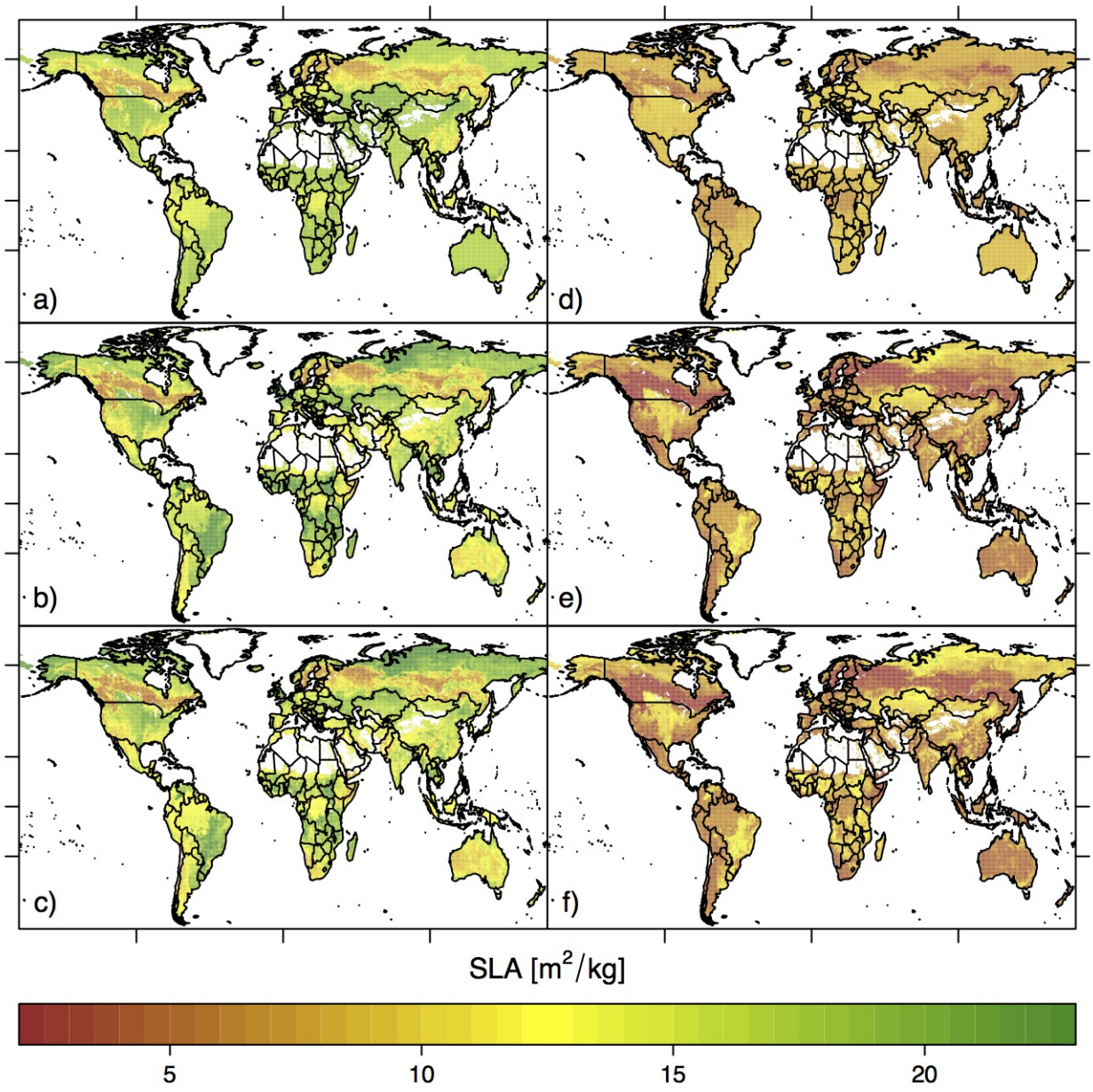

Figure S10: Narrow PFT SLA Models [m^2 kg^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

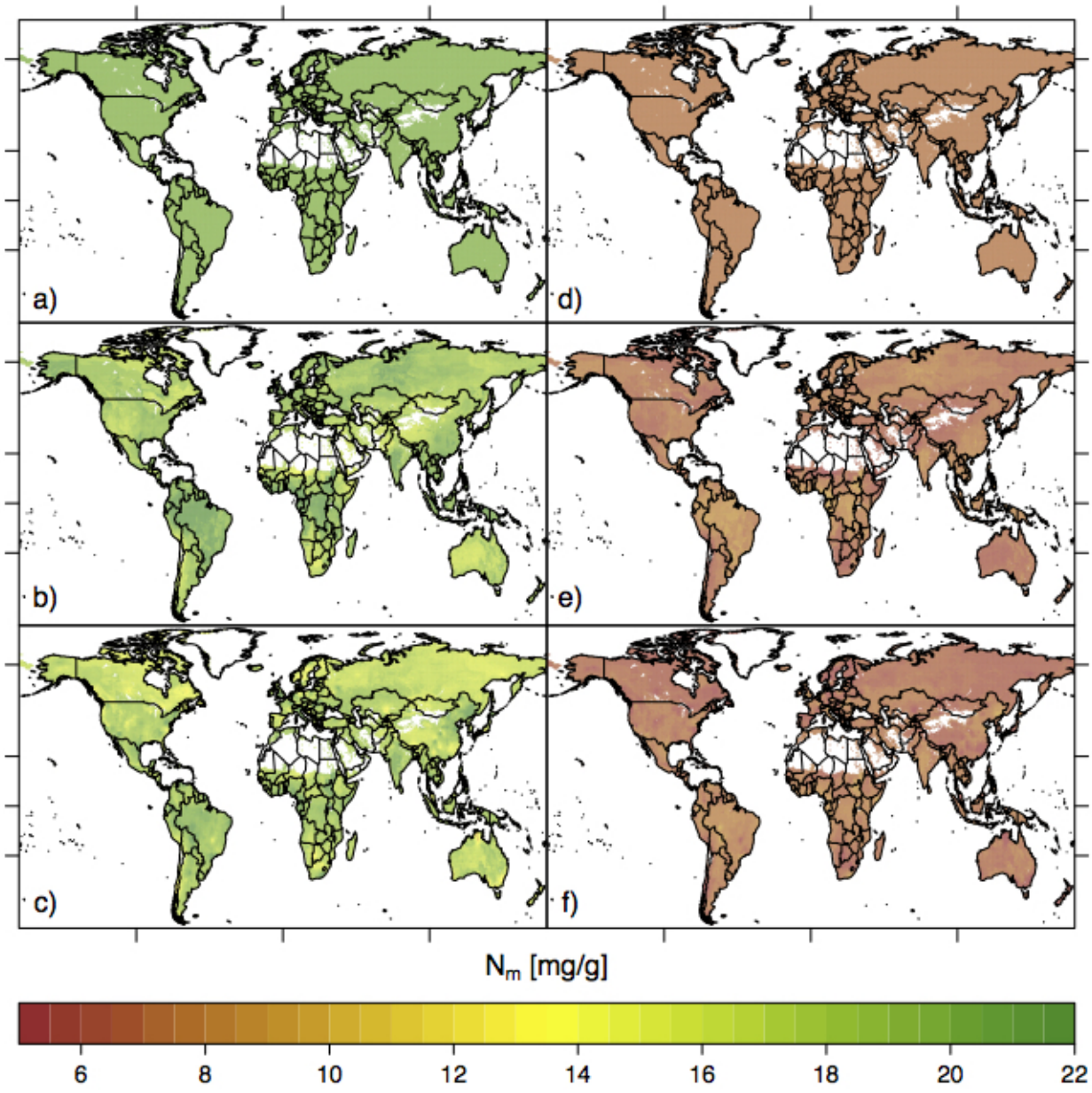

Figure S11: PFT-free N_m Models [mg N g^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

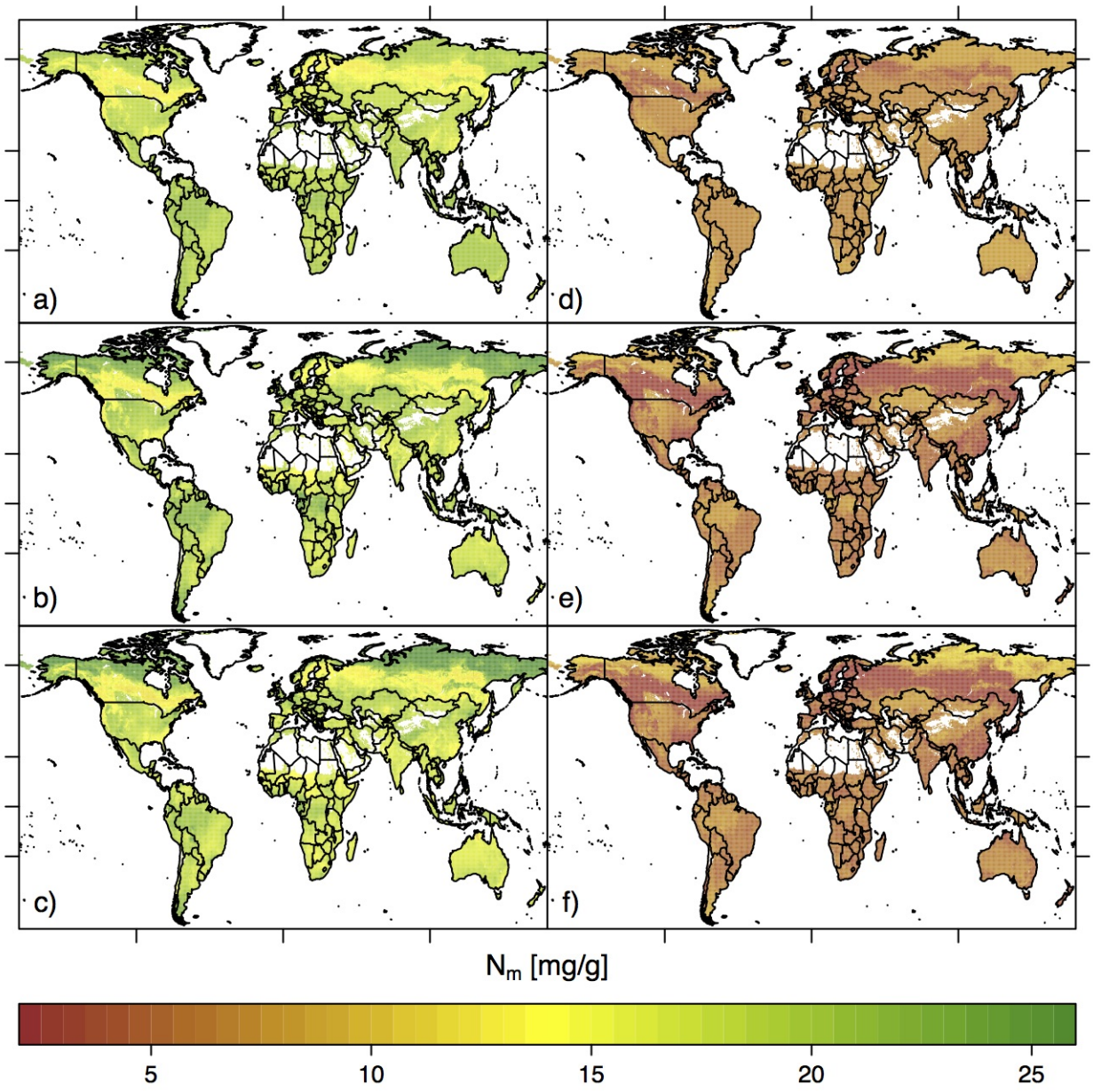

Figure S12: Broad PFT N_m Models [mg N g^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

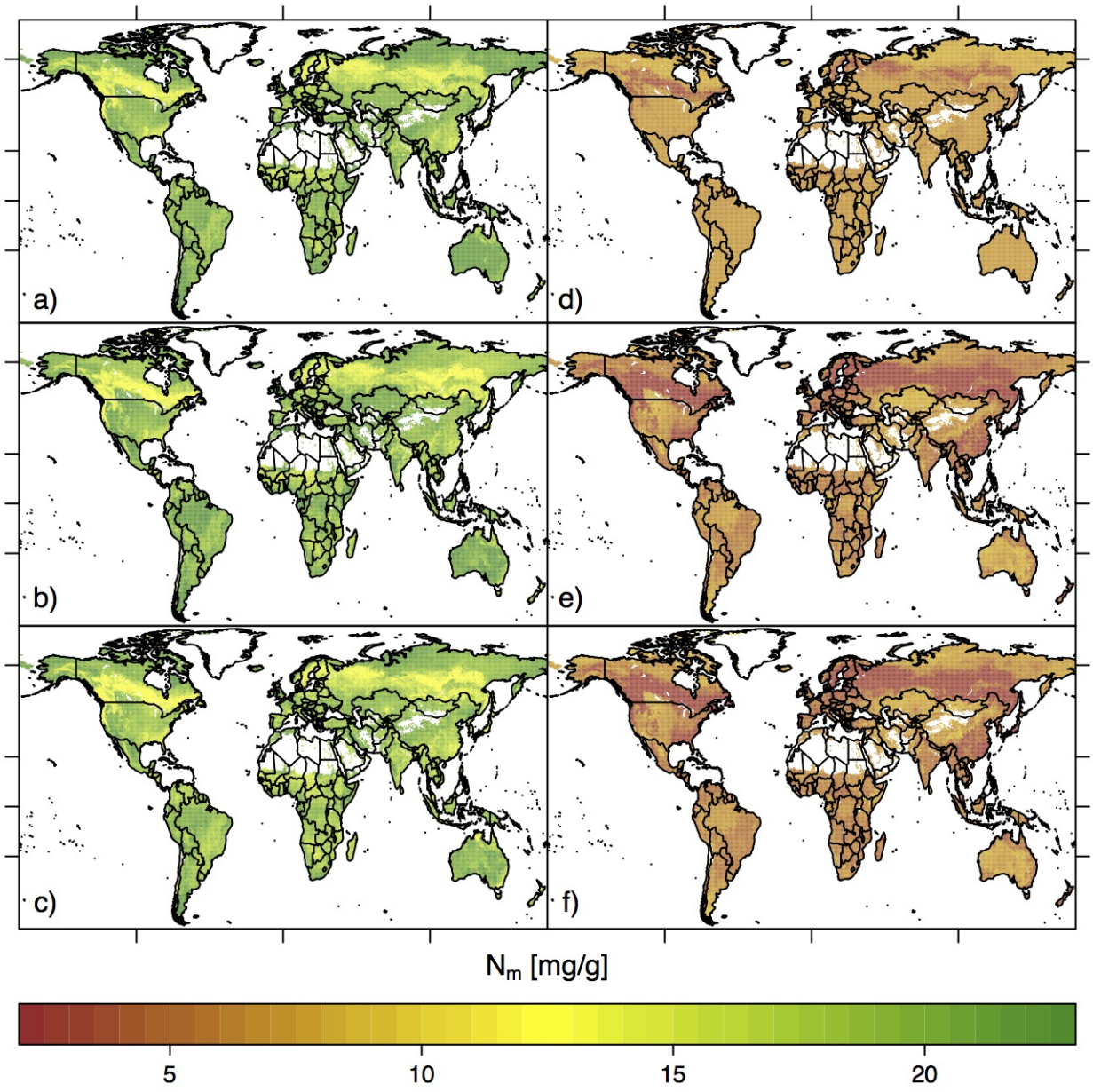

Figure S13: Narrow PFT N_m Models [mg N g^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

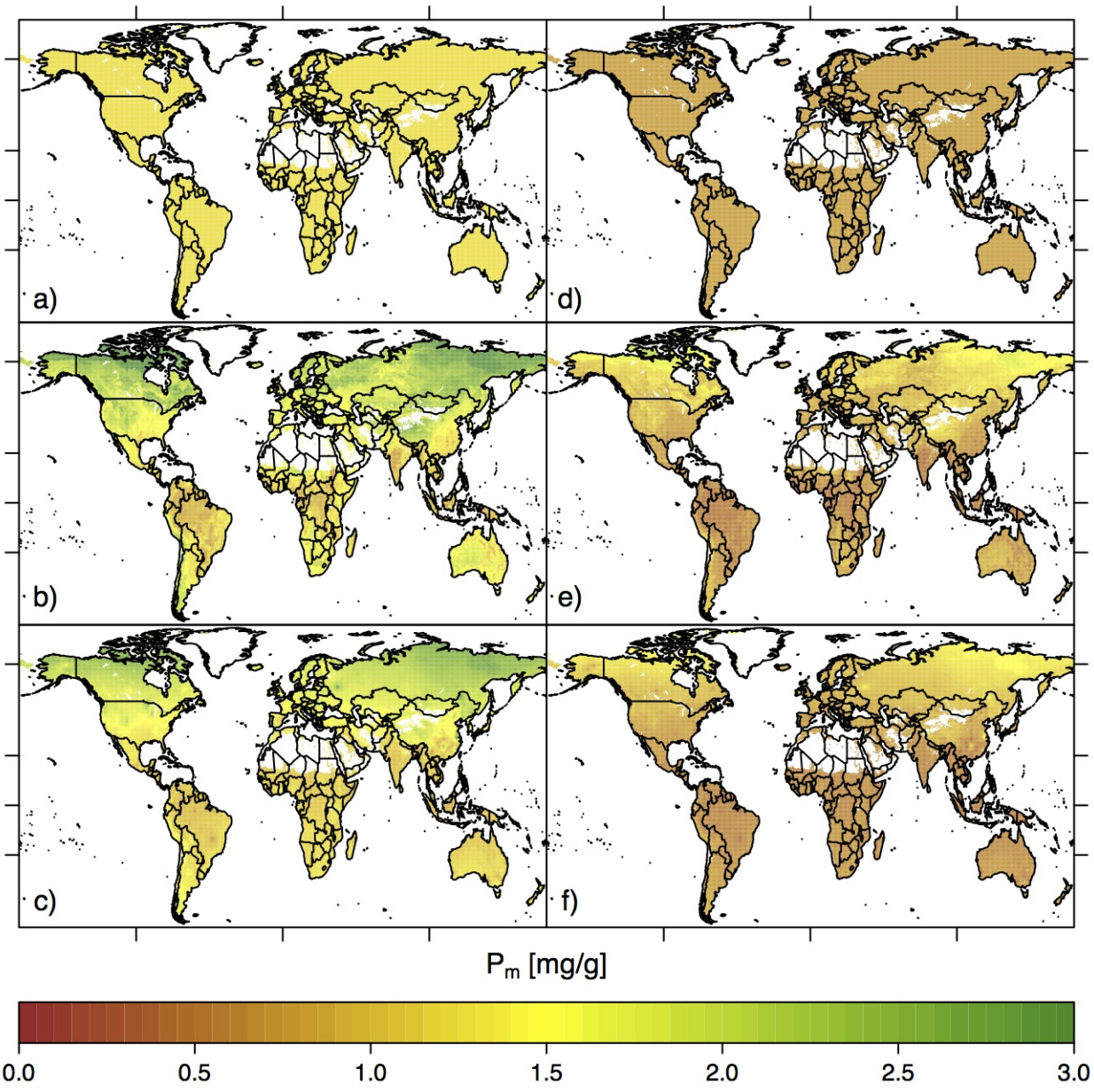

Figure S14: PFT-free P_m Models [mg P g^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

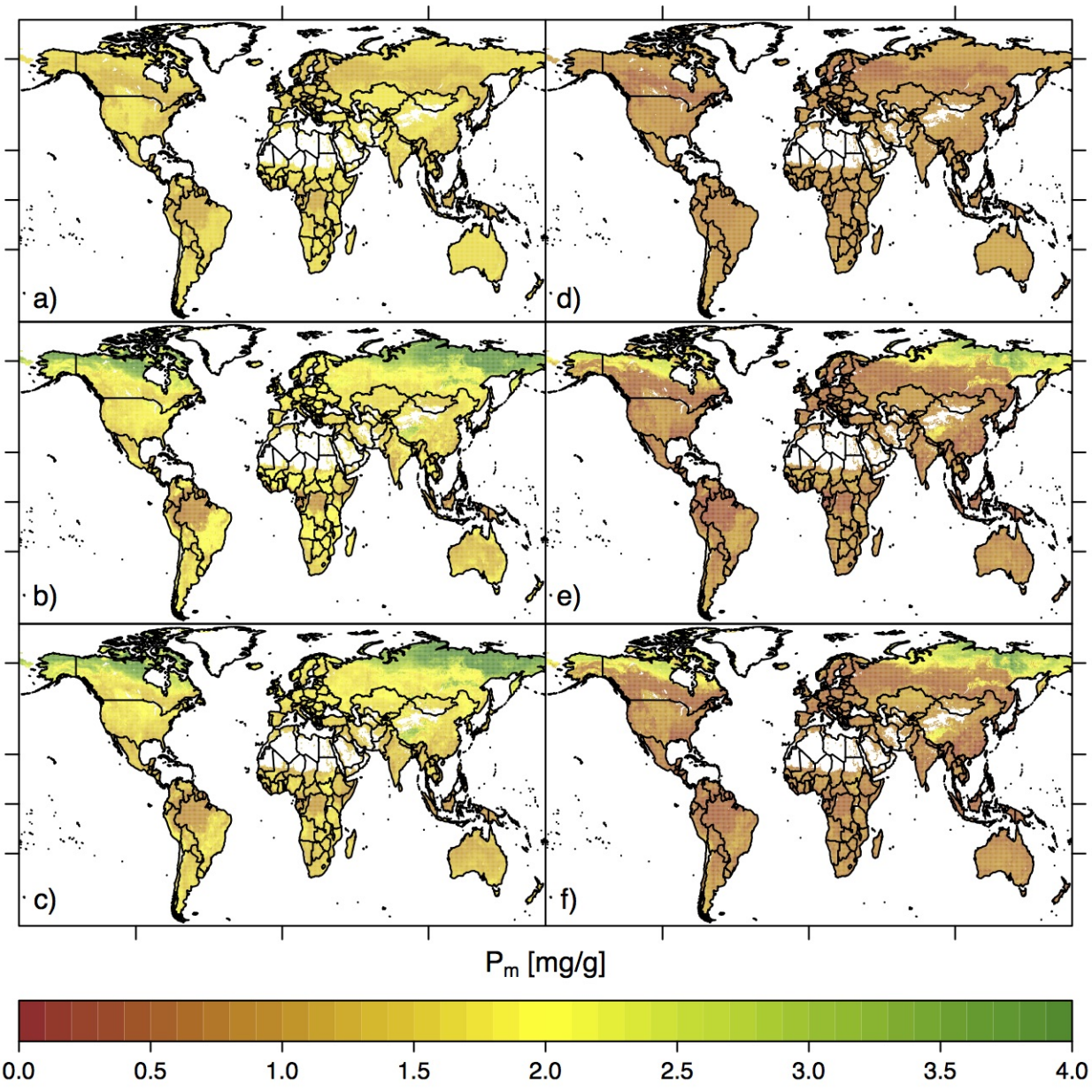

Figure S15: Broad PFT P_m Models [mg P g^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

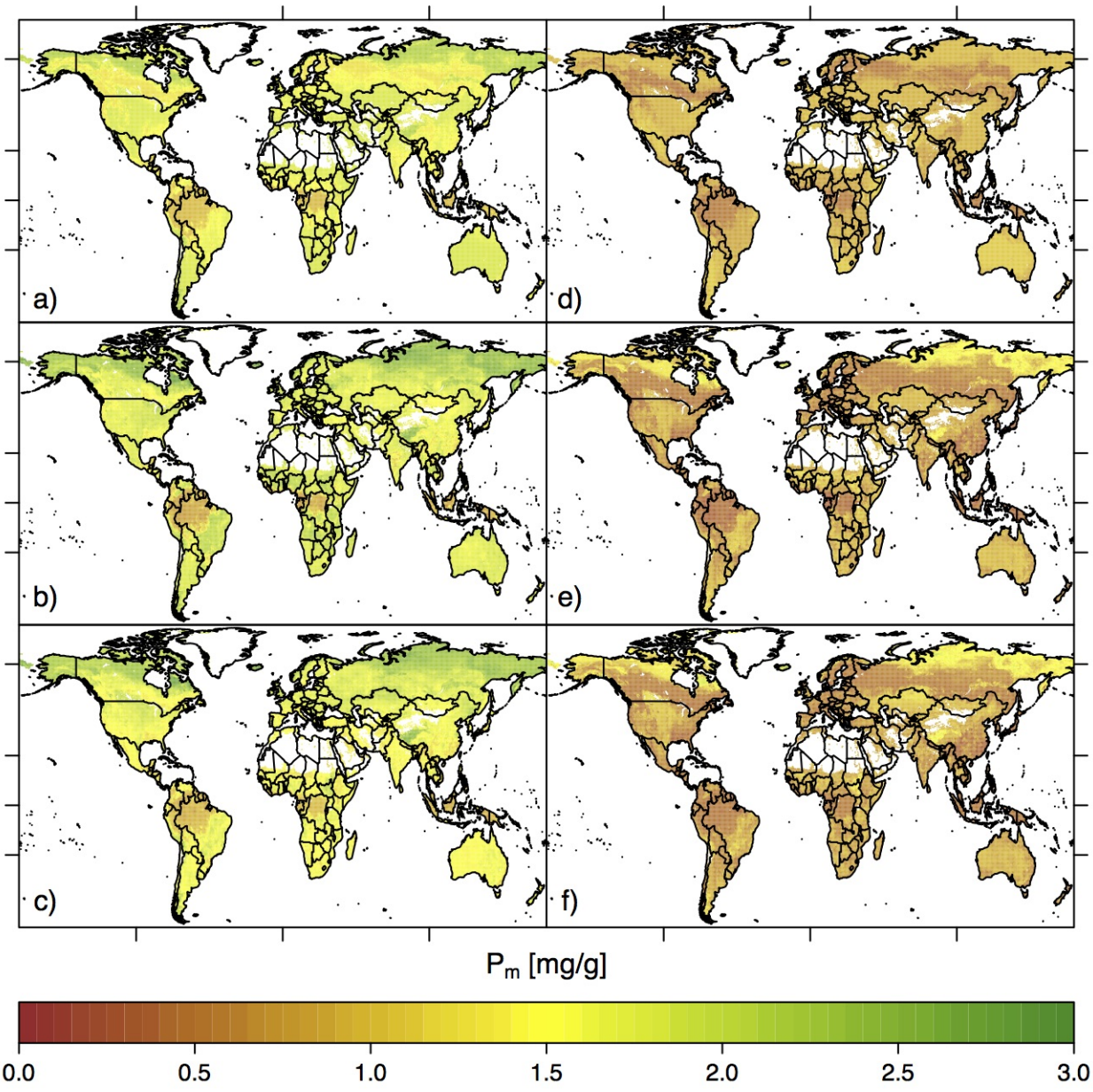

Figure S16: Narrow PFT P_m Models [mg P g^{-1}]. a) categorical mean b) linear mean c) spatial mean d) categorical standard deviation e) linear standard deviation f) spatial standard deviation

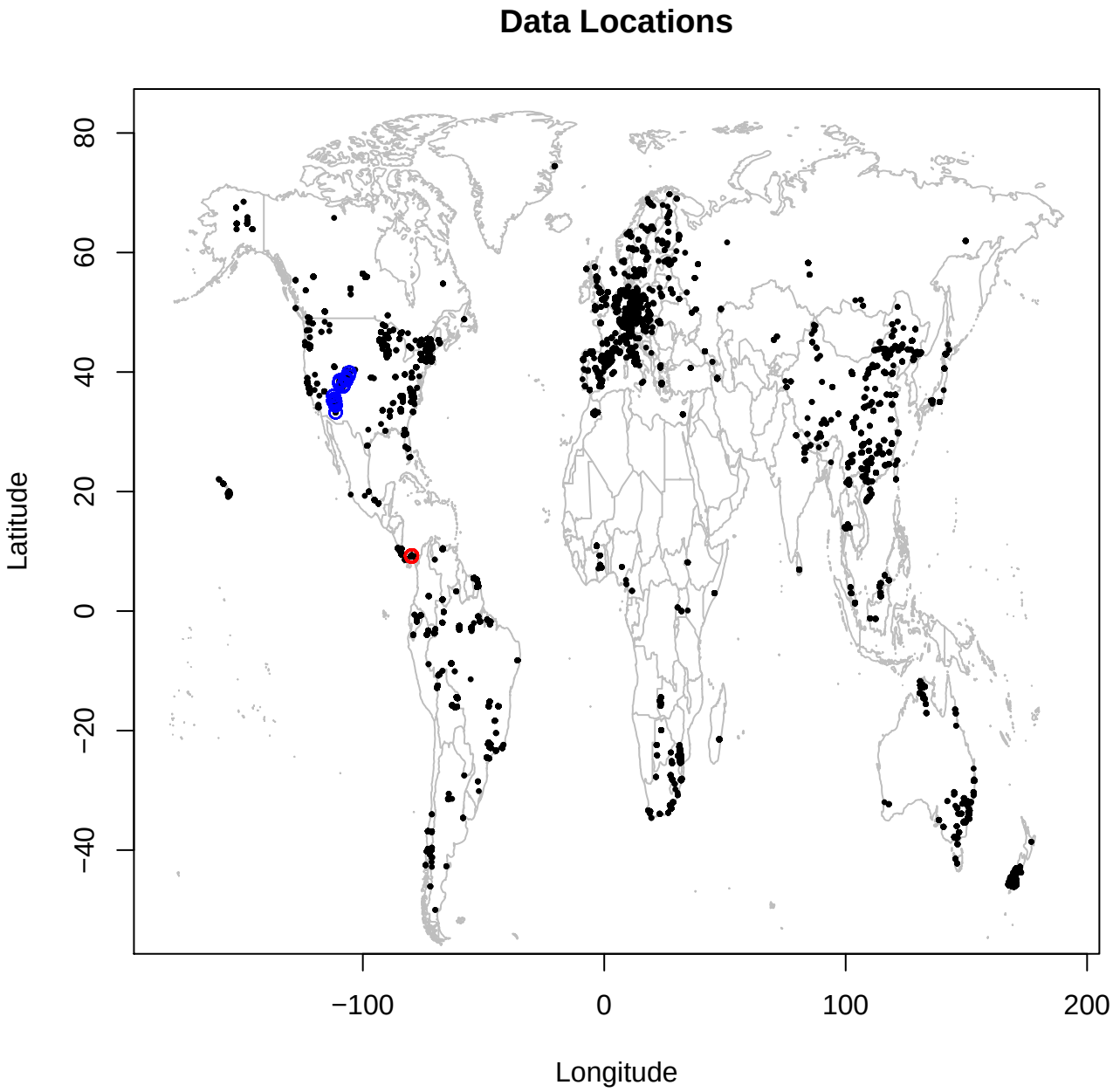

Figure S17: Site Locations: The black points are all locations with data, the red circle is the pixel that includes Barro Colorado Island and blue circles are the composite of sites used in the US Southwest case study.

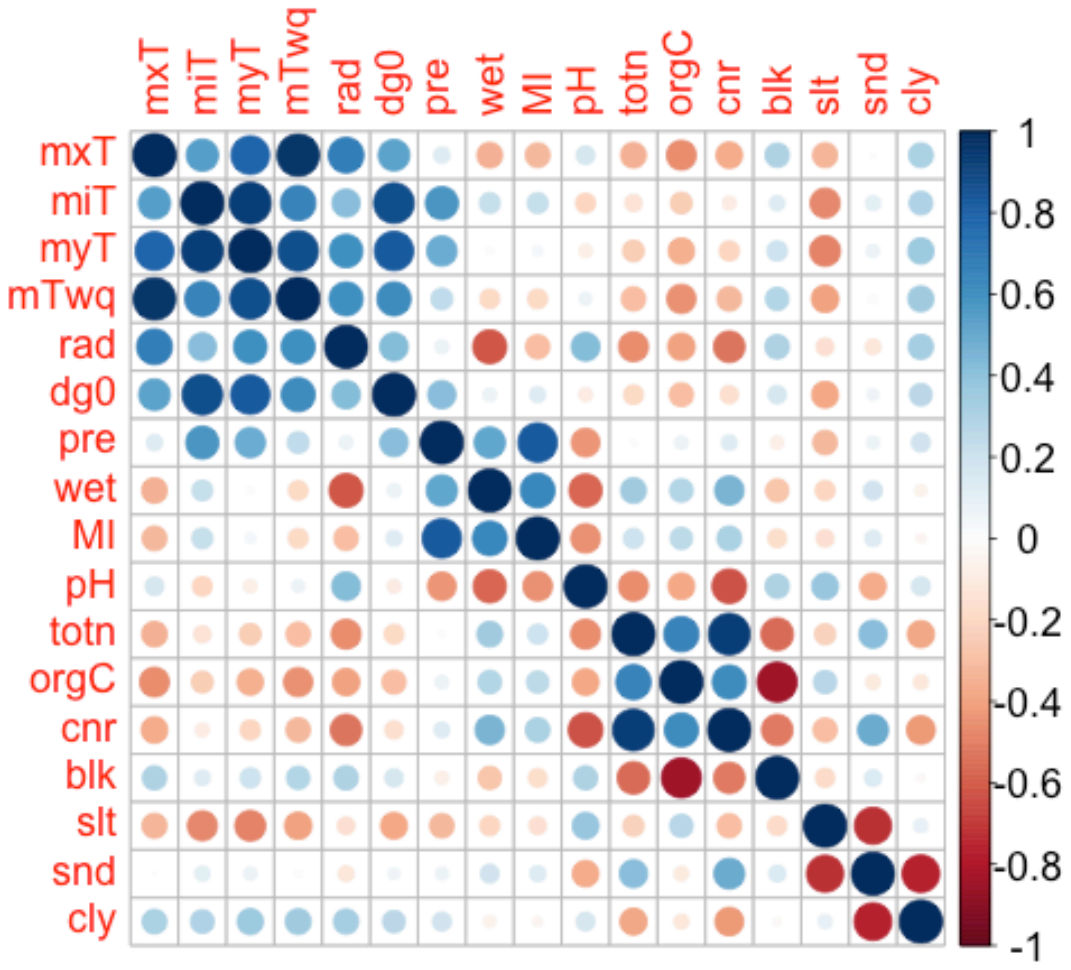

Figure S18: Pairwise correlation for the seventeen environmental predictors. The size of the circles as well as the color indicate the magnitude of the correlation. mxT [C] = Yearly mean monthly maximum temperature; miT [C] = Yearly mean monthly minimum temperature; myT [C] = Yearly mean monthly mean temperature; mTwq [C] = Mean temperature of warmest quarter; rad [w m-2] = Yearly total radiation; dg0 [days] = Number of days above 0C; pre [mm] = Yearly total precipitation; wet [days] = Number of wet days; MI [unitless] = Moisture Index (precipiation/evapotranspiration); pH [pH units] = Percent hydrogen (aqueous); totn [% (mass)] = Percent nitrogen content; orgC [% (mass)] = Percent organic carbon content; cnr [unitless] = Carbon to nitrogen ratio; blk [g cm-3] = Bulk soil density; slt [% (mass)] = Percent silt content; snd [% (mass)] = Percent sand content; cly [% (mass)] = Percent clay content

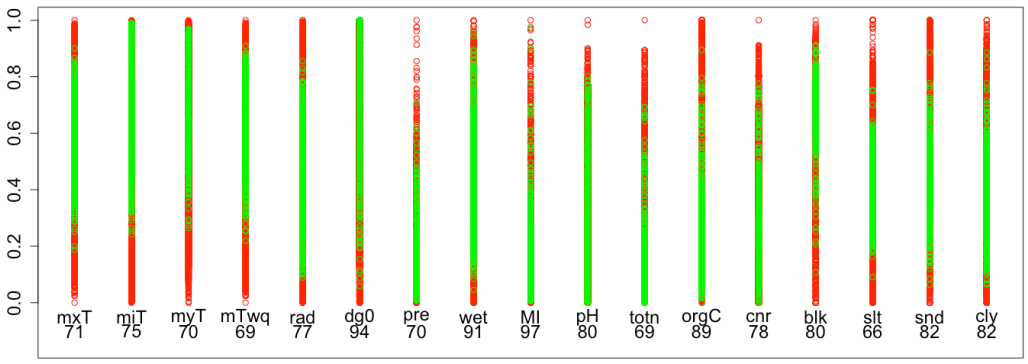

Figure S19: Observed coverage (green) and full coverage (red) of each predictor. The number below each predictor represents the observed to full coverage ratio. mxT [C] = Yearly mean monthly maximum temperature; miT [C] = Yearly mean monthly minimum temperature; myT [C] = Yearly mean monthly mean temperature; mTwq [C] = Mean temperature of warmest quarter; rad [w m-2] = Yearly total radiation; dg0 [days] = Number of days above 0C; pre [mm] = Yearly total precipitation; wet [days] = Number of wet days; MI [unitless] = Moisture Index (precipiation/evapotranspiration); pH [pH units] = Percent hydrogen (aqueous); totn [% (mass)] = Percent nitrogen content; orgC [% (mass)] = Percent organic carbon content; cnr [unitless] = Carbon to nitrogen ratio; blk [g cm-3] = Bulk soil density; slt [% (mass)] = Percent silt content; snd [% (mass)] = Percent sand content; cly [% (mass)] = Percent clay content

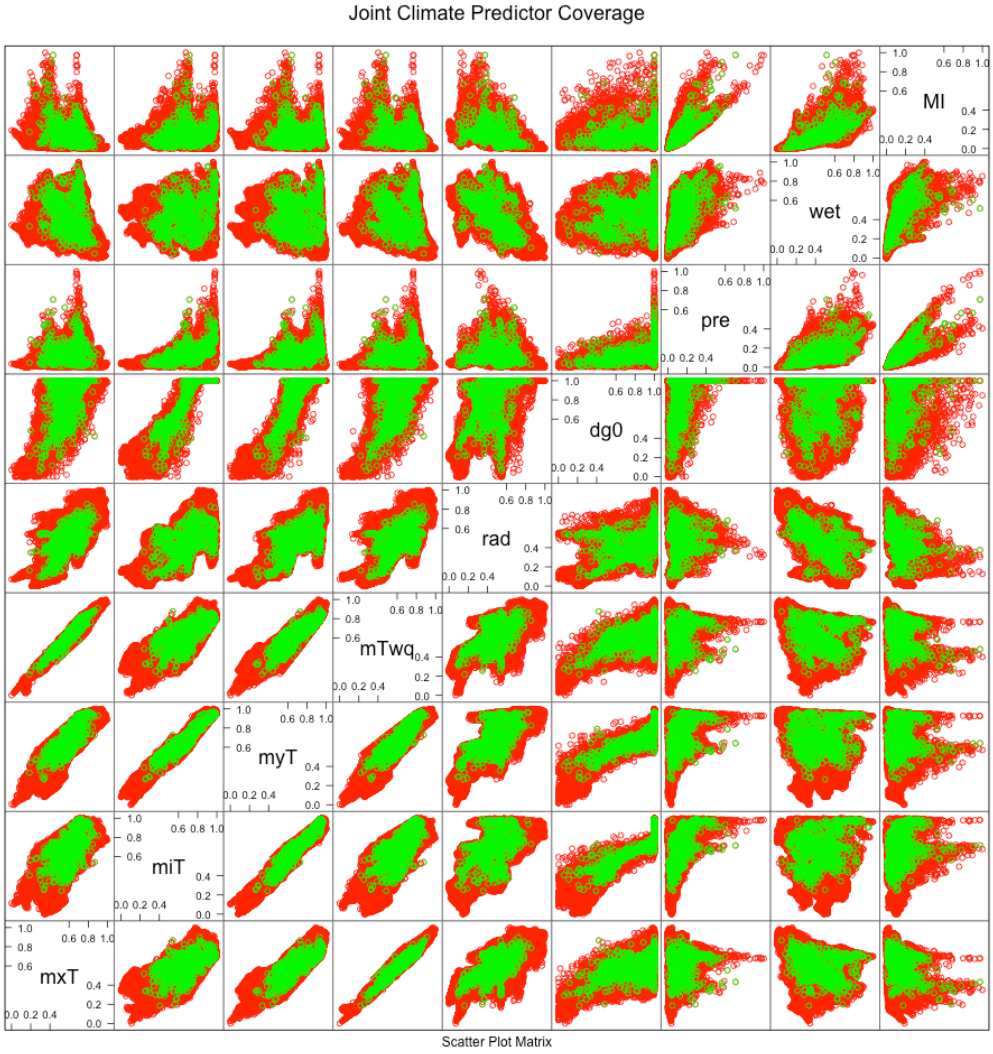

Figure S20: Pairwise predictor coverage for the climate predictors: Green points indicate observed scenarios and red points indicate total range of scenarios. mxT [C] = Yearly mean monthly maximum temperature; miT [C] = Yearly mean monthly minimum temperature; myT [C] = Yearly mean monthly mean temperature; mTwq [C] = Mean temperature of warmest quarter; rad [w m-2] = Yearly total radiation; dg0 [days] = Number of days above 0C; pre [mm] = Yearly total precipitation; wet [days] = Number of wet days; MI [unitless] = Moisture Index (precipiation/evapotranspiration)

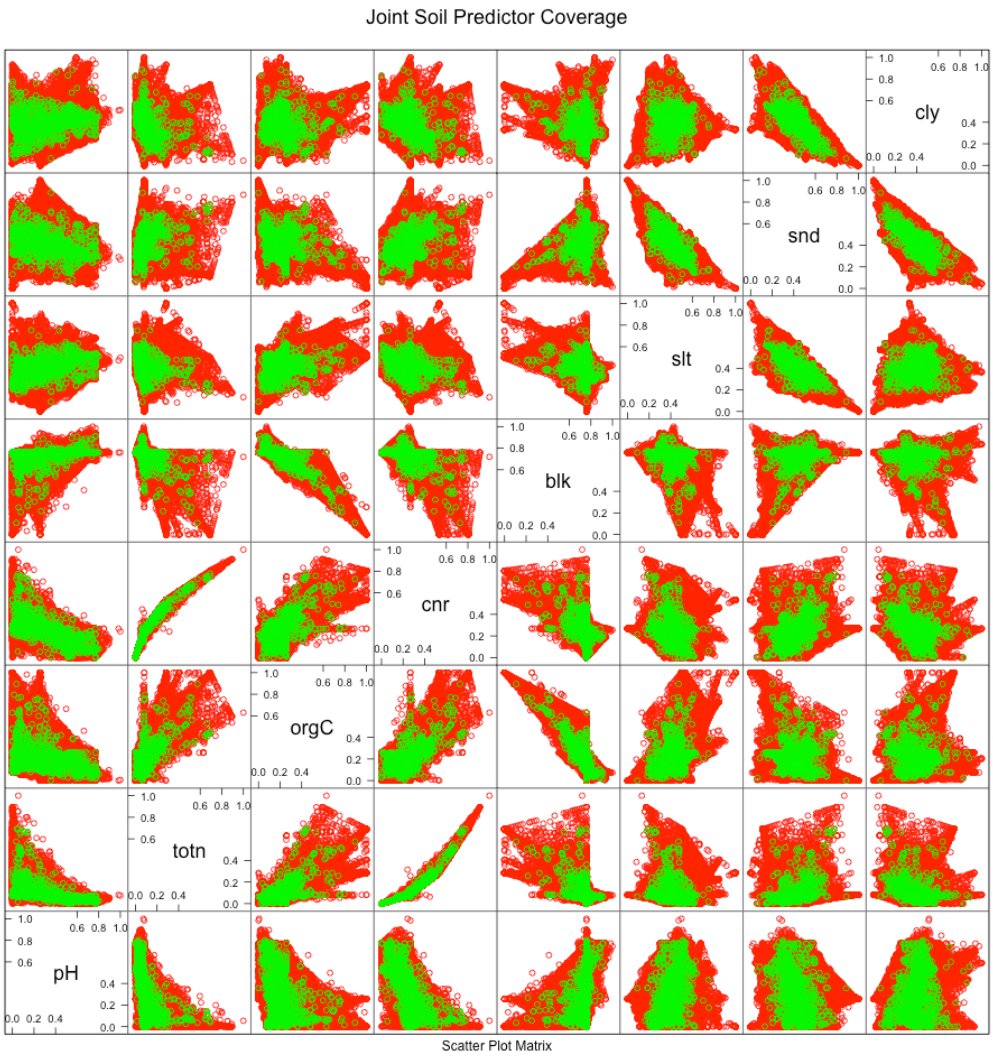

Figure S21: Pairwise predictor coverage for the soil predictors: Green points indicate observed scenarios, red points indicate total range of scenarios. pH [pH units] = Percent hydrogen (aqueous); totn [% (mass)] = Percent nitrogen content; orgC [% (mass)] = Percent organic carbon content; cnr [unitless] = Carbon to nitrogen ratio; blk [g cm-3] = Bulk soil density; slt [% (mass)] = Percent silt content; snd [% (mass)] = Percent sand content; cly [% (mass)] = Percent clay content

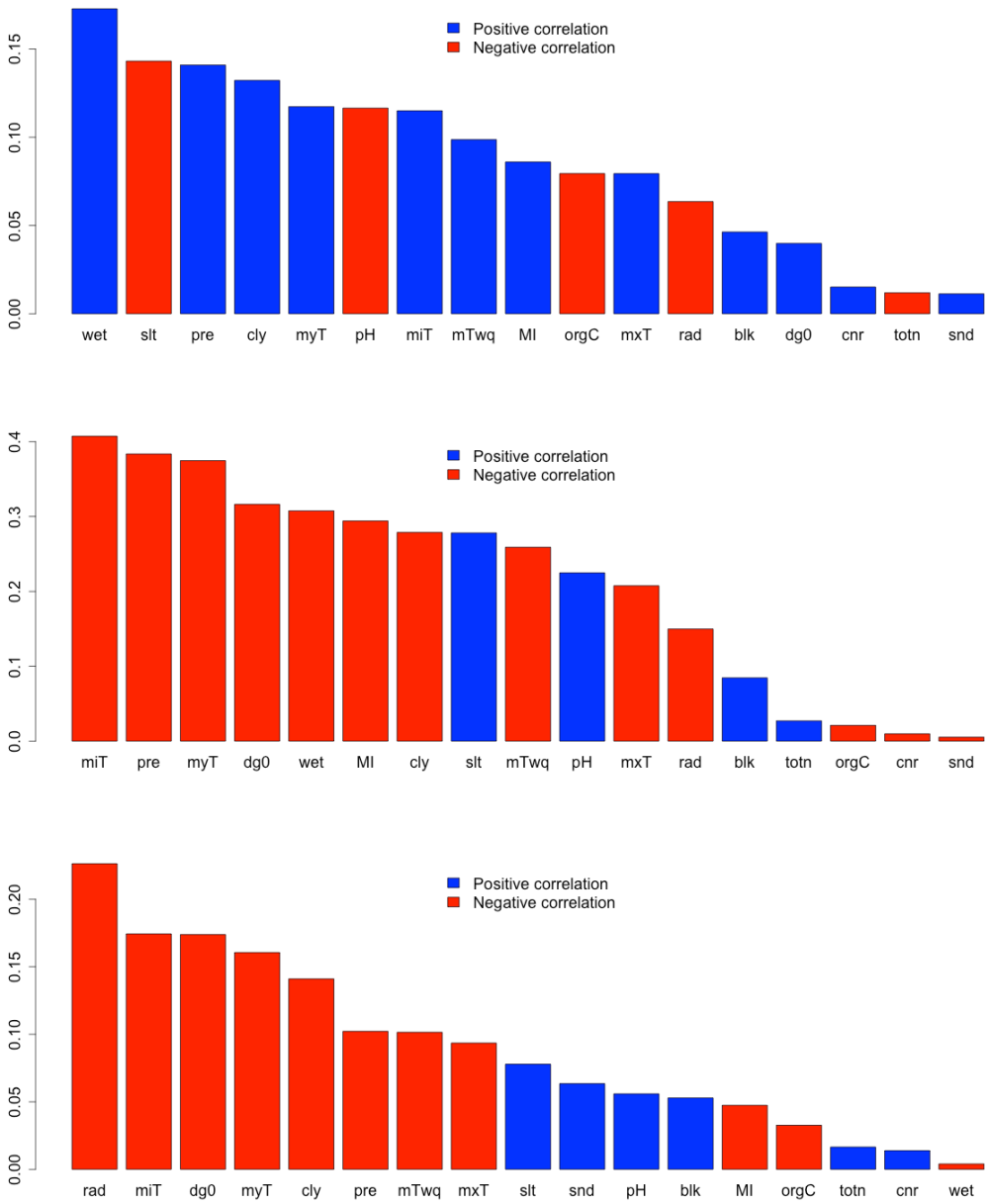

Figure S22: Top: Pairwise absolute correlations between SLA and the seventeen predictors (Red bars indicate positive correlation and blue bars indicate negative correlation). Middle: Pairwise absolute correlations between N_m and the seventeen predictors. Bottom: Pairwise absolute correlations between P_m and the seventeen predictors. mxT [C] = Yearly mean monthly maximum temperature; miT [C] = Yearly mean monthly minimum temperature; myT [C] = Yearly mean monthly mean temperature; mTwq [C] = Mean temperature of warmest quarter; rad [w m-2] = Yearly total radiation; dg0 [days] = Number of days above 0C; pre [mm] = Yearly total precipitation; wet [days] = Number of wet days; MI [unitless] = Moisture Index (precipiation/evapotranspiration); pH [pH units] = Percent hydrogen (aqueous); totn [% (mass)] = Percent nitrogen content; orgC [% (mass)] = Percent organic carbon content; cnr [unitless] = Carbon to nitrogen ratio; blk [g cm-3] = Bulk soil density; slt [% (mass)] = Percent silt content; snd [% (mass)] = Percent sand content; cly [% (mass)] = Percent clay content

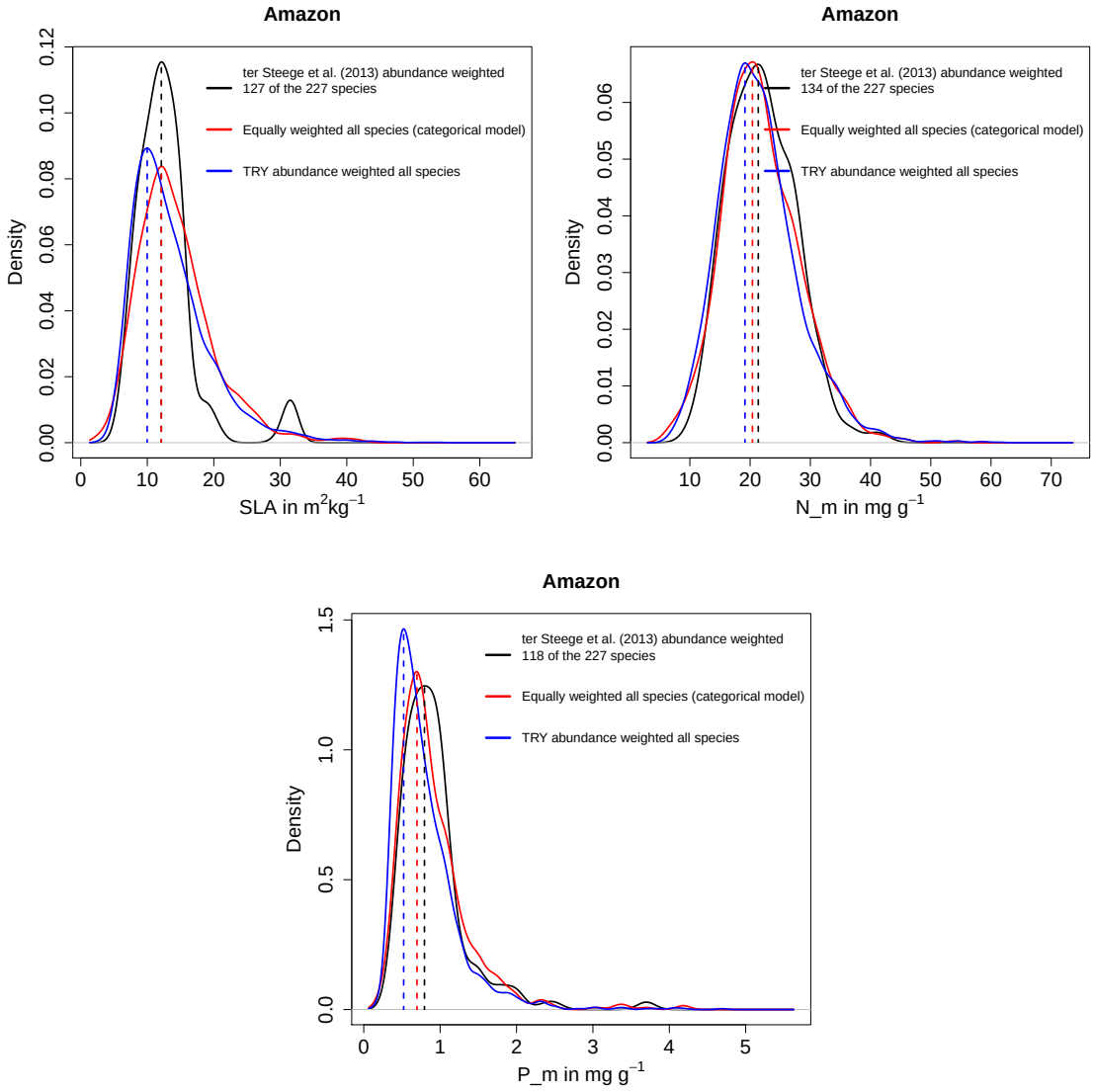

Figure S23: Evaluating the influence of hyperdominance in the Amazon. Each figure corresponds to a trait from top left (clockwise): SLA, N_m, P_m.

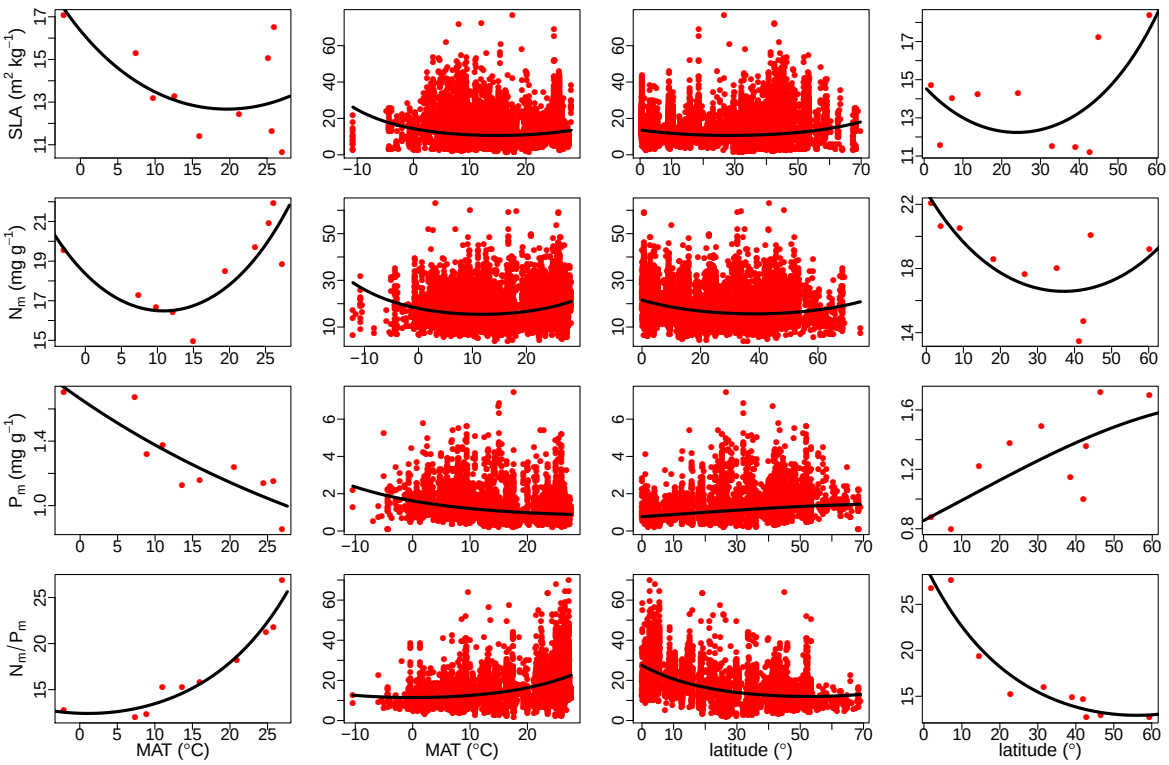

Figure S24: Macro ecology trait-environment relationships. The raw data show different trait-environment relationships between four traits: specific leaf area (SLA, top row), leaf nitrogen (N_m, second row), leaf phosphorus (P_m, third row), and the nitrogen-phosphorus ratio (N_m/P_m, fourth row) and two environment metrics: mean annual temperature (MAT, left two columns) and latitude (right two columns). The outer columns bin the traits by their quantiles while the inner columns are the full data set. Each panel is fit with a quadratic to highlight the broad trait-environment relationship. While phosphorus patterns are consistent with earlier work (Reich and Oleksyn 2004), the pattern for leaf nitrogen may indicate different control mechanisms in hotter and colder climates

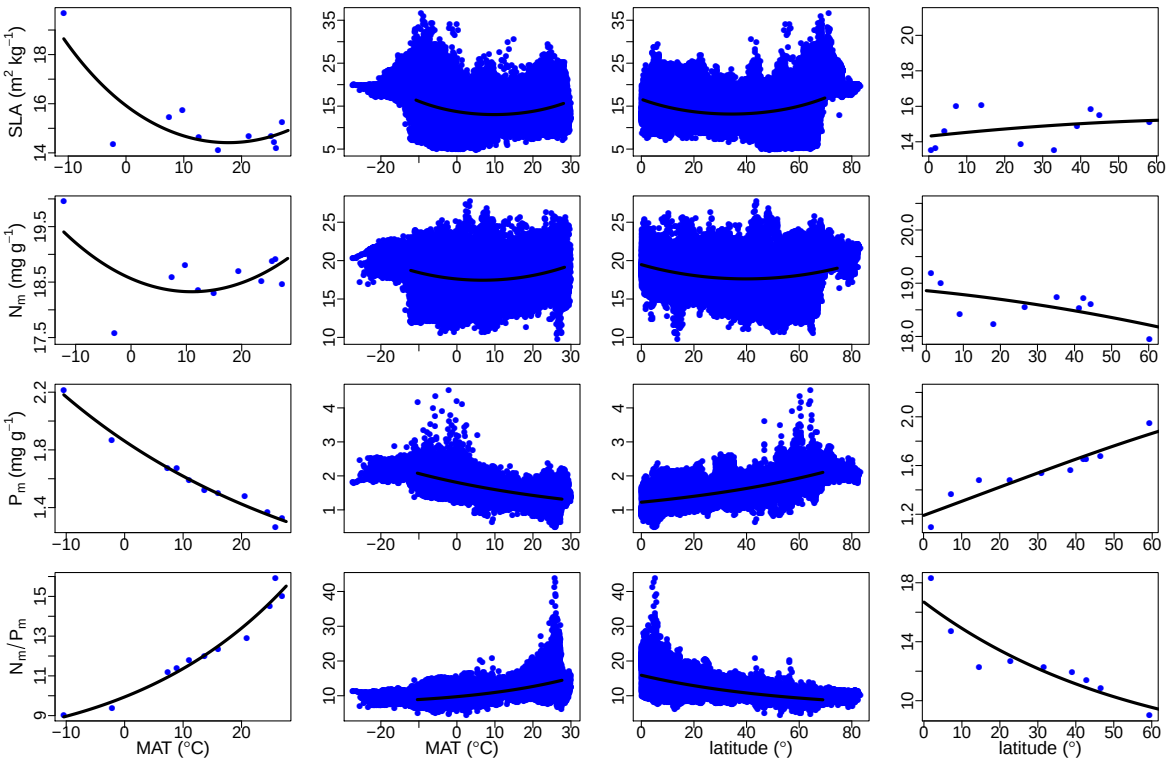

Figure S25: Modeled Trait-Environment Relationships. The spatial model shows similar patterns to the raw data (Figure S24), and is laid out identically: specific leaf area (SLA, top row), leaf nitrogen (N_m, second row), leaf phosphorus (P_m, third row), and the nitrogen-phosphorus ratio (N_m/P_m, fourth row) and two environment metrics: mean annual temperature (MAT, left two columns) and latitude (right two columns). The outer columns bin the traits by their quantiles while the inner columns are the full data set. Each panel is fit with a quadratic to highlight the broad trait-environment relationship. The coherence of these patterns reinforces the suitability of the spatial model for a range of applications.

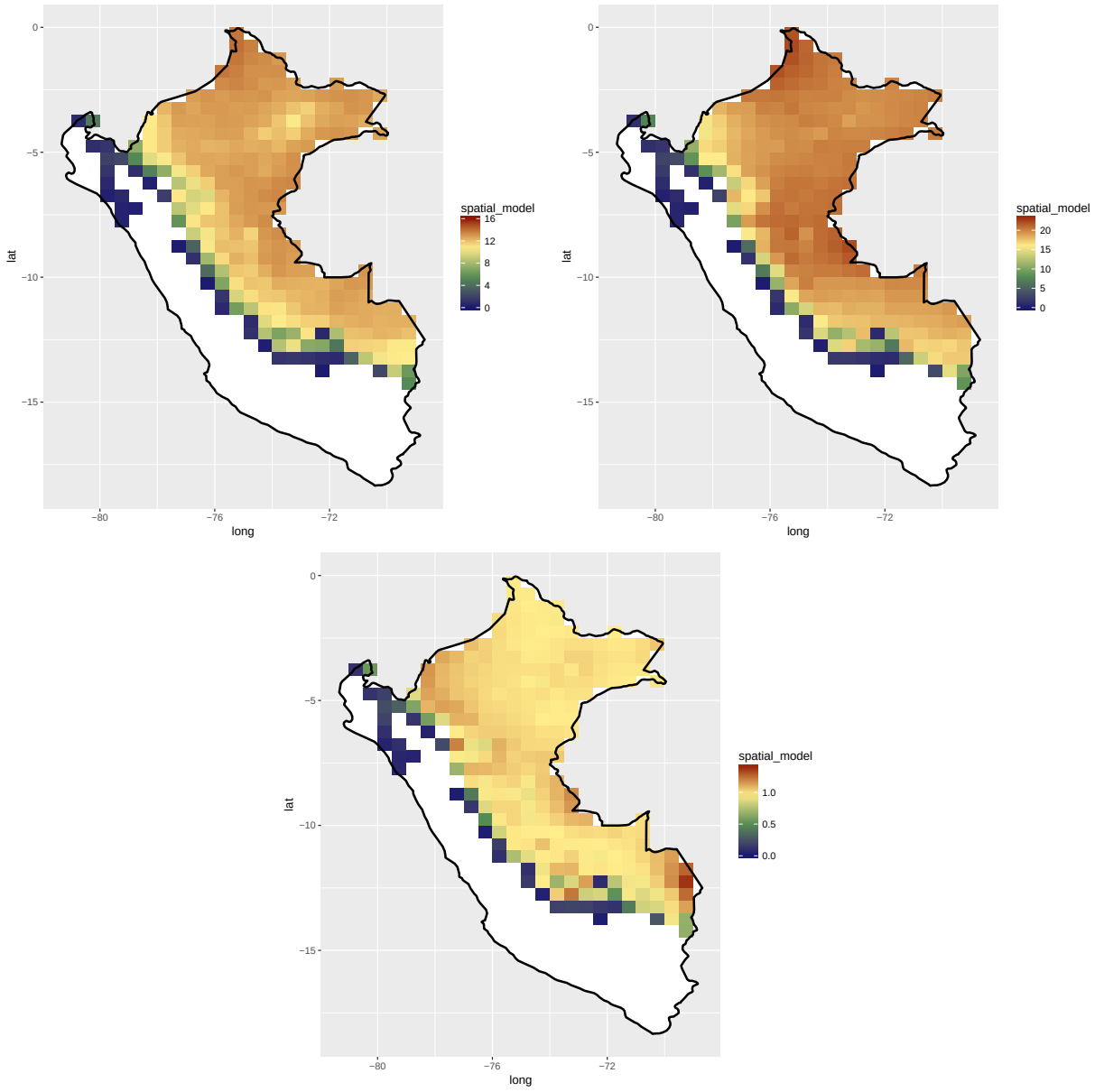

Figure S26: Broadleaf Evergreen Tree estimates in Peru. The close-up maps of Peru show similar patterns, particularly along the east-west elevation gradient as recent detailed remote sensing analyses (Asner et al. 2017). From the upper left and rotating clockwise the traits are: Specific Leaf Area [$m^2 kg^{-1}$], Leaf Nitrogen [mg g^{-1}], and Leaf Phosphorus [mg g^{-1}].

S10 Original TRY References

Ackerly DD, Cornwell WK. **2007**. A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecology Letters* **10**: 135–145.

Adler PB. 2003. A comparison of livestock grazing effects on sagebrush steppe, USA, and Patagonian steppe, Argentina. PhD Thesis, Colorado State University

Adler PB, Milchunas DG, Lauenroth WK, Sala OE, Burke IC. **2004**. Functional traits of graminoids in semi-arid steppes: a test of grazing histories. *Journal of Applied Ecology* **41**: 653–663.

Adriaenssens S. 2012. Dry deposition and canopy exchange for temperate tree species under high nitrogen deposition. PhD thesis, Ghent University, Ghent, Belgium, 209p.

Atkin KO, Schortemeyer M, McFarlane N, Evans RJ. **1999**. The response of fast- and slow-growing Acacia species to elevated atmospheric CO2: an analysis of the underlying components of relative growth rate. *Oecologia* **120**: 544–554.

Atkin OK, Westbeek MHM, Cambridge ML, Lambers H, Pons TL. **1997**. Leaf Respiration in Light and Darkness (A Comparison of Slow- and Fast-Growing Poa Species). *Plant Physiology* **113**: 961–965.

Auger, S. 2012. MSc thesis, Université de Sherbrooke, Sherbrooke (Quebec) L'importance de la variabilité interspécifique des traits fonctionnels par rapport à la variabilité intraspécifique chez les jeunes arbres en forêt mature.

Auger S, Shipley B. **2013**. Inter-specific and intra-specific trait variation along short environmental gradients in an old-growth temperate forest. *Journal of Vegetation Science* **24**: 419–428.

Baker TR, Phillips OL, Laurance WF, Pitman NCA, Almeida S, Arroyo L, DiFiore A, Erwin T, Higuchi N, Killeen TJ, *et al.* 2009. Do species traits determine patterns of wood production in Amazonian forests? *Biogeosciences* **6**: 297–307.

Bakker C, Van Bodegom PM, Nelissen HJM, Ernst WHO, Aerts R. **2006**. Plant responses to rising water tables and nutrient management in calcareous dune slacks. *Plant Ecology* **185**: 19–28.

Bakker C, Rodenburg J, van Bodegom PM. **2005**. Effects of Ca- and Fe-rich Seepage on P Availability and Plant Performance in Calcareous Dune Soils. *Plant and Soil* **275**: 111–122.

Baraloto C, Timothy Paine CE, Patiño S, Bonal D, Hérault B, Chave J. **2010a**. Functional trait variation and sampling strategies in species-rich plant communities. *Functional Ecology* **24**: 208–216.

Baraloto C, Timothy Paine CE, Poorter L, Beauchene J, Bonal D, Domenach A-M, Hérault B, Patiño S, Roggy J-C, Chave J. **2010b**. Decoupled leaf and stem economics in rain forest trees. *Ecology Letters* **13**: 1338–1347.

Beckmann M, Hock M, Bruelheide H, Erfmeier A. **2012**. The role of UV-B radiation in the invasion of Hieracium pilosella—A comparison of German and New Zealand plants. *Environmental and Experimental Botany* **75**: 173–180.

Blonder B, Buzzard V, Simova I, Sloat L, Boyle B, Lipson R, Aguilar-Beaucage B, Andrade A, Barber B, Barnes C, *et al.* 2012. The leaf-area shrinkage effect can bias paleoclimate and

ecology research. *American Journal of Botany* **99**: 1756–1763.
Blonder B, Violle C, Bentley LP, Enquist BJ. **2011**. Venation networks and the origin of the leaf economics spectrum. *Ecology Letters* **14**: 91–100.
Blonder B, Violle C, Enquist BJ. **2013**. Assessing the causes and scales of the leaf economics spectrum using venation networks in Populus tremuloides. *Journal of Ecology* **101**: 981–989.
Blonder, B., Vasseur, F., Violle, C., Shipley, B., Enquist, B., Vile, D. Arabidopsis thaliana rejects theories for the origin of the leaf economics spectrum. (in review, New Phytologist).
van Bodegom PM, Sorrell BK, Oosthoek A, Bakker C, Aerts R. **2008**. SEPARATING THE EFFECTS OF PARTIAL SUBMERGENCE AND SOIL OXYGEN DEMAND ON PLANT PHYSIOLOGY. *Ecology* **89**: 193–204.
Bond-Lamberty B, Wang C, Gower ST. **2002a**. Aboveground and belowground biomass and sapwood area allometric equations for six boreal tree species of northern Manitoba. *Canadian Journal of Forest Research* **32**: 1441–1450.
Bond-Lamberty B, Wang C, Gower ST. **2003**. The use of multiple measurement techniques to refine estimates of conifer needle geometry. *Canadian Journal of Forest Research* **33**: 101–105.
Bond-Lamberty B, Wang C, Gower ST. **2004**. Net primary production and net ecosystem production of a boreal black spruce wildfire chronosequence. *Global Change Biology* **10**: 473–487.
Bond-Lamberty B, Wang C, Gower ST, Norman J. **2002b**. Leaf area dynamics of a boreal black spruce fire chronosequence. *Tree Physiology* **22**: 993–1001.
Brown KA, Flynn DFB, Abram NK, Ingram JC, Johnson SE, Wright P. **2011**. Assessing Natural Resource Use by Forest-Reliant Communities in Madagascar Using Functional Diversity and Functional Redundancy Metrics. *PLoS ONE* **6**: e24107.
Butterfield BJ, Briggs JM. **2011**. Regeneration niche differentiates functional strategies of desert woody plant species. *Oecologia* **165**: 477–487.
Byun C, de Blois S, Brisson J. **2013**. Plant functional group identity and diversity determine biotic resistance to invasion by an exotic grass. *Journal of Ecology* **101**: 128–139.
Campbell C, Atkinson L, Zaragoza-Castells J, Lundmark M, Atkin O, Hurry V. **2007**. Acclimation of photosynthesis and respiration is asynchronous in response to changes in temperature regardless of plant functional group. *New Phytologist* **176**: 375–389.
Campetella G, Botta-Dukát Z, Wellstein C, Canullo R, Gatto S, Chelli S, Mucina L, Bartha S. **2011**. Patterns of plant trait–environment relationships along a forest succession chronosequence. *Agriculture, Ecosystems & Environment* **145**: 38–48.
Carswell FE, Meir P, Wandelli E V, Bonates LCM, Kruijt B, Barbosa EM, Nobre AD, Grace J, Jarvis PG. **2000**. Photosynthetic capacity in a central Amazonian rain forest. *Tree Physiology* **20**: 179–186.
Castro-Díez P, Puyravaud PJ, Cornelissen CJH. **2000**. Leaf structure and anatomy as related to leaf mass per area variation in seedlings of a wide range of woody plant species and types. *Oecologia* **124**: 476–486.
Castro-Díez P, Puyravaud PJ, Cornelissen CJH, Villar-Salvador P. **1998**. Stem anatomy and relative growth rate in seedlings of a wide range of woody plant species and types. *Oecologia* **116**: 57–66.
Cavender-Bares J, Keen A, Miles B. **2006**. Phylogenetic Structure of Floridian Plant Communities Depdends on Taxanomic and Spatial Scale. *Ecology* **87**: S109–S122.

Cavender-Bares J, Sack L, Savage J. **2007**. Atmospheric and soil drought reduce nocturnal conductance in live oaks. *Tree Physiology* **27**: 611–620.
Cerabolini BEL, Brusa G, Ceriani RM, De Andreis R, Luzzaro A, Pierce S. **2010**. Can CSR classification be generally applied outside Britain? *Plant Ecology* **210**: 253–261.
Chen Y, Han W, Tang L, Tang Z, Fang J. **2013**. Leaf nitrogen and phosphorus concentrations of woody plants differ in responses to climate, soil and plant growth form. *Ecography* **36**: 178–184.
Choat B, Jansen S, Brodribb TJ, Cochard H, Delzon S, Bhaskar R, Bucci SJ, Feild TS, Gleason SM, Hacke UG, *et al.* 2012. Global convergence in the vulnerability of forests to drought. *Nature* **491**: 752–755.
Choat B, Sack L, Holbrook NM. **2007**. Diversity of hydraulic traits in nine Cordia species growing in tropical forests with contrasting precipitation. *New Phytologist* **175**: 686–698.
Coomes DA, Heathcote S, Godfrey ER, Shepherd JJ, Sack L. **2008**. Scaling of xylem vessels and veins within the leaves of oak species. *Biology Letters* **4**: 302 LP-306.
Cornelissen JHC. **1996**. An Experimental Comparison of Leaf Decomposition Rates in a Wide Range of Temperate Plant Species and Types. *Journal of Ecology* **84**: 573–582.
Cornelissen CJH. **1999**. A triangular relationship between leaf size and seed size among woody species: allometry, ontogeny, ecology and taxonomy. *Oecologia* **118**: 248–255.
Cornelissen J, Aerts R, Cerabolini B, Werger M, van der Heijden M. **2001**. Carbon cycling traits of plant species are linked with mycorrhizal strategy. *Oecologia* **129**: 611–619.
Cornelissen JHC, Cerabolini B, Castro-Díez P, Villar-Salvador P, Montserrat-Martí G, Puyravaud JP, Maestro M, Werger MJA, Aerts R. **2003**. Functional traits of woody plants: correspondence of species rankings between field adults and laboratory-grown seedlings? *Journal of Vegetation Science* **14**: 311–322.
Cornelissen JHC, Diez PC, Hunt R. **1996**. Seedling Growth, Allocation and Leaf Attributes in a Wide Range of Woody Plant Species and Types. *Journal of Ecology* **84**: 755–765.
Cornelissen JHC, Pérez-Harguindeguy N, Díaz S, Grime JP, Marzano B, Cabido M, Vendramini F, Cerabolini B. **1999**. Leaf structure and defence control litter decomposition rate across species and life forms in regional floras on two continents. *New Phytologist* **143**: 191–200.
Cornelissen JHC, Quested HM, Gwynn-Jones D, Van Logtestijn RSP, De Beus MAH, Kondratchuk A, Callaghan T V., Aerts R. **2004**. Leaf digestibility and litter decomposability are related in a wide range of subarctic plant species and types. *Functional Ecology* **18**: 779–786.
Cornelissen CJH, Werger AMJ, Castro-Díez P, van Rheenen AJW, Rowland PA. **1997**. Foliar nutrients in relation to growth, allocation and leaf traits in seedlings of a wide range of woody plant species and types. *Oecologia* **111**: 460–469.
Cornwell WK, Ackerly DD. **2009**. Community assembly and shifts in plant trait distributions across an environmental gradient in coastal California. *Ecological Monographs* **79**: 109–126.
Cornwell WK, Bhaskar R, Sack L, Cordell S, Lunch CK. **2007**. Adjustment of structure and function of Hawaiian Metrosideros polymorpha at high vs. low precipitation. *Functional Ecology* **21**: 1063–1071.
Cornwell WK, Schwilk DW, Ackerly DD. **2006**. A Trait-Based Test for Habitat Filtering: Convex Hull Volume. *Ecology* **87**: 1465–1471.
Craine JM, Elmore AJ, Aidar MPM, Bustamante M, Dawson TE, Hobbie EA, Kahmen A, Mack MC, McLauchlan KK, Michelsen A, *et al.* 2009. Global patterns of foliar nitrogen isotopes and

their relationships with climate, mycorrhizal fungi, foliar nutrient concentrations, and nitrogen availability. *New Phytologist* **183**: 980–992.
Craine JM, Lee WG, Bond WJ, Williams RJ, Johnson LC. **2005**. Environmental Constraints on a Global Relationship Among Leaf and Root Traits of Grasses. *Ecology* **86**: 12–19.
Craine JM, Nippert JB, Towne EG, Tucker S, Kembel SW, Skibbe A, McLauchlan KK. **2011**. Functional consequences of climate change-induced plant species loss in a tallgrass prairie. *Oecologia* **165**: 1109–1117.
Craine JM, Ocheltree TW, Nippert JB, Towne EG, Skibbe AM, Kembel SW, Fargione JE. **2013**. Global diversity of drought tolerance and grassland climate-change resilience. *Nature Clim. Change* **3**: 63–67.
Craine JM, Towne EG, Ocheltree TW, Nippert JB. **2012**. Community traitscape of foliar nitrogen isotopes reveals N availability patterns in a tallgrass prairie. *Plant and Soil* **356**: 395–403.
Craven D, Braden D, Ashton MS, Berlyn GP, Wishnie M, Dent D. **2007**. Between and within-site comparisons of structural and physiological characteristics and foliar nutrient content of 14 tree species at a wet, fertile site and a dry, infertile site in Panama. *Forest Ecology and Management* **238**: 335–346.
Craven D, Dent D, Braden D, Ashton MS, Berlyn GP, Hall JS. **2011**. Seasonal variability of photosynthetic characteristics influences growth of eight tropical tree species at two sites with contrasting precipitation in Panama. *Forest Ecology and Management* **261**: 1643–1653.
Demey A, Staelens J, Baeten L, Boeckx P, Hermy M, Kattge J, Verheyen K. **2013**. Nutrient input from hemiparasitic litter favors plant species with a fast-growth strategy. *Plant and Soil* **371**: 53–66.
Diaz S, Hodgson JG, Thompson K, Cabido M, Cornelissen JHC, Jalili A, Montserrat-Martí G, Grime JP, Zarrinkamar F, Asri Y, *et al.* 2004. The plant traits that drive ecosystems: Evidence from three continents. *Journal of Vegetation Science* **15**: 295–304.
Domingues TF, Berry JA, Martinelli LA, Ometto JPHB, Ehleringer JR. **2005**. Parameterization of Canopy Structure and Leaf-Level Gas Exchange for an Eastern Amazonian Tropical Rain Forest (Tapajós National Forest, Pará, Brazil). *Earth Interactions* **9**: 1–23.
Domingues FT, Martinelli AL, Ehleringer RJ. **2007**. Ecophysiological traits of plant functional groups in forest and pasture ecosystems from eastern Amazônia, Brazil. *Plant Ecology* **193**: 101–112.
Domingues TF, Meir P, Feldpausch TEDR, Saiz G, Veenendaal EM, Schrodt F, Bird M, Djagbletey G, Hien F, Compaore H, *et al.* 2010. Co-limitation of photosynthetic capacity by nitrogen and phosphorus in West Africa woodlands. *Plant, Cell & Environment* **33**: 959–980.
Dunbar-Co S, Sporck M, Sack L. **2009**. Leaf Trait Diversification and Design in Seven Rare Taxa of the Hawaiian *Plantago* Radiation. *International Journal of Plant Sciences* **170**: 61–75.
E. Medlyn B, Jarvis PG. **1999**. Design and use of a database of model parameters from elevated [CO2] experiments. *Ecological Modelling* **124**: 69–83.
Fitter AH, Peat HJ. **1994**. The Ecological Flora Database. *Journal of Ecology* **82**: 415–425.
Fonseca CR, Overton JM, Collins B, Westoby M. **2000**. Shifts in trait-combinations along rainfall and phosphorus gradients. *Journal of Ecology* **88**: 964–977.
Fortunel C, Garnier E, Joffre R, Kazakou E, Quested H, Grigulis K, Lavorel S, Ansquer P, Castro H, Cruz P, *et al.* 2009. Leaf traits capture the effects of land use changes and climate on litter decomposability of grasslands across Europe. *Ecology* **90**: 598–611.

Frenette-Dussault C, Shipley B, Léger J-F, Meziane D, Hingrat Y. **2012**. Functional structure of an arid steppe plant community reveals similarities with Grime's C-S-R theory. *Journal of Vegetation Science* **23**: 208–222.
Freschet GT, Cornelissen JHC, van Logtestijn RSP, Aerts R. **2010a**. Substantial nutrient resorption from leaves, stems and roots in a subarctic flora: what is the link with other resource economics traits? *New Phytologist* **186**: 879–889.
Freschet GT, Cornelissen JHC, Van Logtestijn RSP, Aerts R. **2010b**. Evidence of the 'plant economics spectrum' in a subarctic flora. *Journal of Ecology* **98**: 362–373.
Fyllas NM, Patiño S, Baker TR, Bielefeld Nardoto G, Martinelli LA, Quesada CA, Paiva R, Schwarz M, Horna V, Mercado LM, *et al.* 2009. Basin-wide variations in foliar properties of Amazonian forest: phylogeny, soils and climate. *Biogeosciences* **6**: 2677–2708.
Garnier E, Lavorel S, Ansquer P, Castro H, Cruz P, Dolezal J, Eriksson O, Fortunel C, Freitas H, Golodets C, *et al.* 2007. Assessing the Effects of Land-use Change on Plant Traits, Communities and Ecosystem Functioning in Grasslands: A Standardized Methodology and Lessons from an Application to 11 European Sites. *Annals of Botany* **99**: 967–985.
Givnish TJ, Montgomery RA, Goldstein G. **2004**. Adaptive radiation of photosynthetic physiology in the Hawaiian lobeliads: light regimes, static light responses, and whole-plant compensation points. *American Journal of Botany* **91**: 228–246.
Gutiérrez AG, Huth A. **2012**. Successional stages of primary temperate rainforests of Chiloé Island, Chile. *Perspectives in Plant Ecology, Evolution and Systematics* **14**: 243–256.
Guy AL, Mischkolz JM, Lamb EG. **2012**. Limited effects of simulated acidic deposition on seedling survivorship and root morphology of endemic plant taxa of the Athabasca Sand Dunes in well-watered greenhouse trials. *Botany* **91**: 176–181.
Han W, Chen Y, Zhao F-J, Tang L, Jiang R, Zhang F. **2012**. Floral, climatic and soil pH controls on leaf ash content in China's terrestrial plants. *Global Ecology and Biogeography* **21**: 376–382.
Han W, Fang J, Guo D, Zhang Y. **2005**. Leaf nitrogen and phosphorus stoichiometry across 753 terrestrial plant species in China. *The New phytologist* **168**: 377–385.
Hao G-Y, Sack L, Wang A-Y, Cao K-F, Goldstein G. **2010**. Differentiation of leaf water flux and drought tolerance traits in hemiepiphytic and non-hemiepiphytic Ficus tree species. *Functional Ecology* **24**: 731–740.
He J-S, Wang L, Flynn DFB, Wang X, Ma W, Fang J. **2008**. Leaf nitrogen:phosphorus stoichiometry across Chinese grassland biomes. *Oecologia* **155**: 301–310.
He J-S, Wang Z, Wang X, Schmid B, Zuo W, Zhou M, Zheng C, Wang M, Fang J. **2006**. A test of the generality of leaf trait relationships on the Tibetan Plateau. *New Phytologist* **170**: 835–848.
Hickler, T. 1999. Plant functional types and community characteristics along environmental gradients on Öland's Great Alvar (Sweden) Masters Thesis, University of Lund, Sweden.
Hoof J, Sack L, Webb DT, Nilsen ET. **2008**. Contrasting Structure and Function of Pubescent and Glabrous Varieties of Hawaiian Metrosideros polymorpha (Myrtaceae) at High Elevation. *Biotropica* **40**: 113–118.
Kattge J, Knorr W, Raddatz T, Wirth C. **2009**. Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology* **15**: 976–991.
Kazakou E, Vile D, Shipley B, Gallet C, Garnier E. **2006**. Co-variations in litter decomposition, leaf traits and plant growth in species from a Mediterranean old-field succession. *Functional*

Ecology **20**: 21–30.
Kerkoff A., Fagan WF, James J. Elser, Brian J. Enquist. **2006**. Phylogenetic and Growth Form Variation in the Scaling of Nitrogen and Phosphorus in the Seed Plants. *The American Naturalist* **168**: E103–E122.
Kichenin E, Wardle DA, Peltzer DA, Morse CW, Freschet GT. **2013**. Contrasting effects of plant inter- and intraspecific variation on community-level trait measures along an environmental gradient. *Functional Ecology* **27**: 1254–1261.
Kleyer M, Bekker RM, Knevel IC, Bakker JP, Thompson K, Sonnenschein M, Poschlod P, Van Groenendael JM, Klimeš L, Klimešová J, *et al.* 2008. The LEDA Traitbase: a database of life-history traits of the Northwest European flora. *Journal of Ecology* **96**: 1266–1274.
Kraft NJB, Valencia R, Ackerly DD. **2008**. Functional Traits and Niche-Based Tree Community Assembly in an Amazonian Forest. *Science* **322**: 580 LP-582.
Kurokawa H, Nakashizuka T. **2008**. Leaf Herbivory and Decomposability in a Malaysian Tropical Rain Forest. *Ecology* **89**: 2645–2656.
Laughlin DC, Fulé PZ, Huffman DW, Crouse J, Laliberté E. **2011**. Climatic constraints on trait-based forest assembly. *Journal of Ecology* **99**: 1489–1499.
Laughlin DC, Leppert JJ, Moore MM, Sieg CH. **2010**. A multi-trait test of the leaf-height-seed plant strategy scheme with 133 species from a pine forest flora. *Functional Ecology* **24**: 493–501.
Louault F, Pillar VD, Aufrère J, Garnier E, Soussana J-F. **2005**. Plant traits and functional types in response to reduced disturbance in a semi-natural grassland. *Journal of Vegetation Science* **16**: 151–160.
LOVEYS BR, ATKINSON LJ, SHERLOCK DJ, ROBERTS RL, FITTER AH, ATKIN OK. **2003**. Thermal acclimation of leaf and root respiration: an investigation comparing inherently fast- and slow-growing plant species. *Global Change Biology* **9**: 895–910.
Manning P, Newington JE, Robson HR, Saunders M, Eggers T, Bradford MA, Bardgett RD, Bonkowski M, Ellis RJ, Gange AC, *et al.* 2006. Decoupling the direct and indirect effects of nitrogen deposition on ecosystem function. *Ecology Letters* **9**: 1015–1024.
Markesteijn L, Poorter L, Paz H, Sack L, Bongers F. **2011**. Ecological differentiation in xylem cavitation resistance is associated with stem and leaf structural traits. *Plant, Cell & Environment* **34**: 137–148.
Martin RE, Asner GP, Sack L. **2007**. Genetic variation in leaf pigment, optical and photosynthetic function among diverse phenotypes of Metrosideros polymorpha grown in a common garden. *Oecologia* **151**: 387–400.
McDonald PG, Fonseca CR, Overton JM, Westoby M. **2003**. Leaf-size divergence along rainfall and soil-nutrient gradients: is the method of size reduction common among clades? *Functional Ecology* **17**: 50–57.
Mckenna MF, Shipley B. **1999**. Interacting determinants of interspecific relative growth: Empirical patterns and a theoretical explanation. *Écoscience* **6**: 286–296.
Medlyn BE, Badeck F-W, De Pury DGG, Barton CVM, Broadmeadow M, Ceulemans R, De Angelis P, Forstreuter M, Jach ME, Kellomäki S, *et al.* 1999. Effects of elevated [CO2] on photosynthesis in European forest species: a meta-analysis of model parameters . *Plant, Cell & Environment* **22**: 1475–1495.
Medlyn BE, Barton CVM, Broadmeadow MSJ, Ceulemans R, De Angelis P, Forstreuter M,

Freeman M, Jackson SB, Kellomäki S, Laitat E, *et al.* 2001. Stomatal conductance of forest species after long-term exposure to elevated CO2 concentration: a synthesis. *New Phytologist* **149**: 247–264.
Meir P, Kruijt B, Broadmeadow M, Barbosa E, Kull O, Carswell F, Nobre A, Jarvis PG. **2002**. Acclimation of photosynthetic capacity to irradiance in tree canopies in relation to leaf nitrogen concentration and leaf mass per unit area. *Plant, Cell & Environment* **25**: 343–357.
Meir P, Levy PE, Grace J, Jarvis PG. **2007**. Photosynthetic parameters from two contrasting woody vegetation types in West Africa. *Plant Ecology* **192**: 277–287.
Meng T-T, Wang H, Harrison SP, Prentice IC, Ni J, Wang G. **2015**. Responses of leaf traits to climatic gradients: adaptive variation versus compositional shifts. *Biogeosciences* **12**: 5339–5352.
Messier J, McGill BJ, Lechowicz MJ. **2010**. How do traits vary across ecological scales? A case for trait-based ecology. *Ecology Letters* **13**: 838–848.
Meziane D, Shipley B. **1999**. Interacting components of interspecific relative growth rate: constancy and change under differing conditions of light and nutrient supply. *Functional Ecology* **13**: 611–622.
Meziane D, Shipley B. **1999**. Interacting determinants of specific leaf area in 22 herbaceous species: effects of irradiance and nutrient availability. *Plant, Cell & Environment* **22**: 447–459.
Milla R, Reich PB. **2011**. Multi-trait interactions, not phylogeny, fine-tune leaf size reduction with increasing altitude. *Annals of Botany* **107**: 455–465.
Minden V, Andratschke S, Spalke J, Timmermann H, Kleyer M. **2012**. Plant trait–environment relationships in salt marshes: Deviations from predictions by ecological concepts. *Perspectives in Plant Ecology, Evolution and Systematics* **14**: 183–192.
Minden V, Kleyer M. **2011**. Testing the effect–response framework: key response and effect traits determining above-ground biomass of salt marshes. *Journal of Vegetation Science* **22**: 387–401.
Mischkolz, J. M. 2013. Selecting and evaluating native forage mixtures for the mixed grass prairie. Univeristy of Saskatchewan, Saskatoon, SK.
Nakahashi CD, Frole K, Sack L. **2005**. Bacterial Leaf Nodule Symbiosis in Ardisia (Myrsinaceae): Does it Contribute to Seedling Growth Capacity? *Plant biol (Stuttg)* **7**: 495–500.
Niinemets Ü. **2001**. GLOBAL-SCALE CLIMATIC CONTROLS OF LEAF DRY MASS PER AREA, DENSITY, AND THICKNESS IN TREES AND SHRUBS. *Ecology* **82**: 453–469.
Niinemets Ü. **1999**. Research review. Components of leaf dry mass per area – thickness and density – alter leaf photosynthetic capacity in reverse directions in woody plants. *New Phytologist* **144**: 35–47.
Ogaya R, Penuelas J. **2006**. Contrasting foliar responses to drought in Quercus ilex and Phillyrea latifolia. *Biologia Plantarum* **50**: 373–382.
Ogaya R, Peñuelas J. **2003**. Comparative field study of Quercus ilex and Phillyrea latifolia: photosynthetic response to experimental drought conditions. *Environmental and Experimental Botany* **50**: 137–148.
Ogaya R, Peñuelas J. **2007**. Tree growth, mortality, and above-ground biomass accumulation in a holm oak forest under a five-year experimental field drought. *Plant Ecology* **189**: 291–299.
Ogaya R, Peñuelas J. **2008**. Changes in leaf δ13C and δ15N for three Mediterranean tree species in relation to soil water availability. *Acta Oecologica* **34**: 331–338.

Ordoñez JC, van Bodegom PM, Witte J-PM, Bartholomeus RP, van Dobben HF, Aerts R. **2010**. Leaf habit and woodiness regulate different leaf economy traits at a given nutrient supply. *Ecology* **91**: 3218–3228.
Ordoñez JC, Van Bodegom PM, Witte JPM, Wright IJ, Reich PB, Aerts R. **2009**. A global study of relationships between leaf traits, climate and soil measures of nutrient fertility. *Global Ecology and Biogeography* **18**: 137–149.
Pahl AT, Kollmann J, Mayer A, Haider S. **2013**. No evidence for local adaptation in an invasive alien plant: field and greenhouse experiments tracing a colonization sequence. *Annals of Botany* **112**: 1921–1930.
Pakeman RJ, Garnier E, Lavorel S, Ansquer P, Castro H, Cruz P, Doležal J, Eriksson O, Freitas H, Golodets C, *et al.* 2008. Impact of abundance weighting on the response of seed traits to climate and land use. *Journal of Ecology* **96**: 355–366.
Pakeman RJ, Lepš J, Kleyer M, Lavorel S, Garnier E, consortium the V. **2009**. Relative climatic, edaphic and management controls of plant functional trait signatures. *Journal of Vegetation Science* **20**: 148–159.
Patiño S, Lloyd J, Paiva R, Baker TR, Quesada CA, Mercado LM, Schmerler J, Schwarz M, Santos AJB, Aguilar A, *et al.* 2009. Branch xylem density variations across the Amazon Basin. *Biogeosciences* **6**: 545–568.
Peco B, de Pablos I, Traba J, Levassor C. **2005**. The effect of grazing abandonment on species composition and functional traits: the case of dehesa grasslands. *Basic and Applied Ecology* **6**: 175–183.
Peñuelas J, Sardans J, Llusià J, Owen SM, Carnicer J, Giambelluca TW, Rezende EL, Waite M, Niinemets Ü. **2010**. Faster returns on 'leaf economics' and different biogeochemical niche in invasive compared with native plant species. *Global Change Biology* **16**: 2171–2185.
Peñuelas J, Sardans J, Llusia J, Owen SM, Silva J, Niinemets Ü. **2010**. Higher Allocation to Low Cost Chemical Defenses in Invasive Species of Hawaii. *Journal of Chemical Ecology* **36**: 1255–1270.
Pierce S, Brusa G, Sartori M, Cerabolini BEL. **2012**. Combined use of leaf size and economics traits allows direct comparison of hydrophyte and terrestrial herbaceous adaptive strategies. *Annals of Botany* .
Pierce S, Brusa G, Vagge I, Cerabolini BEL. **2013**. Allocating CSR plant functional types: the use of leaf economics and size traits to classify woody and herbaceous vascular plants. *Functional Ecology* **27**: 1002–1010.
Pierce S, Ceriani RM, DE Andreis R, Luzzaro A, Cerabolini B. **2007**. The leaf economics spectrum of Poaceae reflects variation in survival strategies. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology* **141**: 337–343.
Pierce S, Luzzaro A, Caccianiga M, Ceriani RM, Cerabolini B. **2007**. Disturbance is the principal α-scale filter determining niche differentiation, coexistence and biodiversity in an alpine community. *Journal of Ecology* **95**: 698–706.
Pillar VD, Sosinski EE. **2003**. An improved method for searching plant functional types by numerical analysis. *Journal of Vegetation Science* **14**: 323–332.
Poorter H, Niinemets U, Poorter L, Wright IJ, Villar R. **2009**. Causes and consequences of variation in leaf mass per area (LMA):a meta-analysis. *New Phytologist* **182**: 565–588.
Powers JS, Tiffin P. **2010**. Plant functional type classifications in tropical dry forests in Costa

Rica: leaf habit versus taxonomic approaches. *Functional Ecology* **24**: 927–936.
Prentice IC, Meng T, Wang H, Harrison SP, Ni J, Wang G. **2011**. Evidence of a universal scaling relationship for leaf CO2 drawdown along an aridity gradient. *New Phytologist* **190**: 169–180.
Preston KA, Cornwell WK, DeNoyer JL. **2006**. Wood density and vessel traits as distinct correlates of ecological strategy in 51 California coast range angiosperms. *New Phytologist* **170**: 807–818.
Price CA, Enquist BJ. **2007**. Scaling Mass and Morphology in Leaves: An Extension of the WBE Model. *Ecology* **88**: 1132–1141.
Price CA, Enquist BJ, Savage VM. **2007**. A general model for allometric covariation in botanical form and function. *Proceedings of the National Academy of Sciences* **104**: 13204–13209.
Pyankov VI, Kondratchuk A V, Shipley B. **1999**. Leaf structure and specific leaf mass: the alpine desert plants of the Eastern Pamirs, Tadjikistan. *New Phytologist* **143**: 131–142.
Quero JL, Villar R, Marañón T, Zamora R, Vega D, Sack L. **2008**. Relating leaf photosynthetic rate to whole-plant growth: drought and shade effects on seedlings of four Quercus species. *Functional Plant Biology* **35**: 725–737.
Quested HM, Cornelissen JHC, Press MC, Callaghan T V, Aerts R, Trosien F, Riemann P, Gwynn-Jones D, Kondratchuk A, Jonasson SE. **2003**. Decomposition of Sub-Arctic Plants with Differing Nitrogen Economies: A Functional Role for Hemiparasites. *Ecology* **84**: 3209–3221.
Reich PB, Oleksyn J, Wright IJ. **2009**. Leaf phosphorus influences the photosynthesis–nitrogen relation: a cross-biome analysis of 314 species. *Oecologia* **160**: 207–212.
Reich PB, Tjoelker MG, Pregitzer KS, Wright IJ, Oleksyn J, Machado J-L. **2008**. Scaling of respiration to nitrogen in leaves, stems and roots of higher land plants. *Ecology Letters* **11**: 793–801.
Sack L. **2004**. Responses of temperate woody seedlings to shade and drought: do trade-offs limit potential niche differentiation? *Oikos* **107**: 110–127.
Sack L, Cowan PD, Jaikumar N, Holbrook NM. **2003**. The 'hydrology' of leaves: co-ordination of structure and function in temperate woody species. *Plant, Cell & Environment* **26**: 1343–1356.
Sack L, Frole K. **2006**. Leaf Structural Diversity is Related to Hydraulic Capacity in Tropical Rain Forest Trees. *Ecology* **87**: 483–491.
Sack L, Melcher PJ, Liu WH, Middleton E, Pardee T. **2006**. How strong is intracanopy leaf plasticity in temperate deciduous trees? *American Journal of Botany* **93**: 829–839.
Sack L, Tyree MT, Holbrook NM. **2005**. Leaf hydraulic architecture correlates with regeneration irradiance in tropical rainforest trees. *New Phytologist* **167**: 403–413.
Sandel B, Corbin JD, Krupa M. **2011**. Using plant functional traits to guide restoration: A case study in California coastal grassland. *Ecosphere* **2**: 1–16.
Sardans J, Peñuelas J, Ogaya R. **2008a**. Drought-induced changes in C and N stoichiometry in a Quercus ilex Mediterranean forest. *Forest Science* **54**: 513–522.
Sardans J, Peñuelas J, Prieto P, Estiarte M. **2008b**. Changes in Ca, Fe, Mg, Mo, Na, and S content in a Mediterranean shrubland under warming and drought. *Journal of Geophysical Research: Biogeosciences* **113**: n/a-n/a.
Scherer-Lorenzen M, Schulze E-D, Don A, Schumacher J, Weller E. **2007**. Exploring the functional significance of forest diversity: A new long-term experiment with temperate tree species (BIOTREE). *Perspectives in Plant Ecology, Evolution and Systematics* **9**: 53–70.
Scoffoni C, Pou A, Aasamaa K, SACK L. **2008**. The rapid light response of leaf hydraulic

conductance: new evidence from two experimental methods. *Plant, Cell & Environment* **31**: 1803–1812.
Shipley B. **1989**. The Use of Above-Ground Maximum Relative Growth Rate as an Accurate Predictor of Whole-Plant Maximum Relative Growth Rate. *Functional Ecology* **3**: 771–775.
Shipley B. **1995**. Structured Interspecific Determinants of Specific Leaf Area in 34 Species of Herbaceous Angiosperms. *Functional Ecology* **9**: 312–319.
Shipley B. **2002**. Trade-offs between net assimilation rate and specific leaf area in determining relative growth rate: relationship with daily irradiance. *Functional Ecology* **16**: 682–689.
Shipley B, Lechowicz MJ. **2000**. The functional co-ordination of leaf morphology, nitrogen concentration, and gas exchange in40 wetland species. *Écoscience* **7**: 183–194.
Shipley B, Parent M. **1991**. Germination Responses of 64 Wetland Species in Relation to Seed Size, Minimum Time to Reproduction and Seedling Relative Growth Rate. *Functional Ecology* **5**: 111–118.
Spasojevic, M.J. & Suding, K.N. 2012. Inferring community assembly mechanisms from functional diversity patterns: the importance of multiple assembly processes. *J. Ecol.*, 100, 652-661.
Swaine, E. K. 2007. Ecological and evolutionary drivers of plant community assembly in a Bornean rain forest. PhD Thesis, University of Aberdeen, Aberdeen.
Tucker SS, Craine JM, Nippert JB. **2011**. Physiological drought tolerance and the structuring of tallgrass prairie assemblages. *Ecosphere* **2**: 1–19.
Vergutz L, Manzoni S, Porporato A, Novais RF, Jackson RB. **2012**. Global resorption efficiencies and concentrations of carbon and nutrients in leaves of terrestrial plants. *Ecological Monographs* **82**: 205–220.
Vergutz L, Manzoni S, Porporato A, Novais RF, Jackson RB. **2012**. A Global Database of Carbon and Nutrient Concentrations of Green and Senesced Leaves.
Vile, D. 2005. Significations fonctionnelle et ecologique des traits des especes vegetales: exemple dans une succession post-cultural mediterraneenne et generalisations, PHD Thesis.
de Vries FT, Bardgett RD. **2016**. Plant community controls on short-term ecosystem nitrogen retention. *New Phytologist* **210**: 861–874.
Waite M, Sack L. **2010**. How does moss photosynthesis relate to leaf and canopy structure? Trait relationships for 10 Hawaiian species of contrasting light habitats. *New Phytologist* **185**: 156–172.
Williams M, Shimabukuro YE, Rasetter EB. **2012**. LBA-ECO CD-09 Soil and Vegetation Characteristics, Tapajos National Forest, Brazil.
Willis CG, Halina M, Lehman C, Reich PB, Keen A, McCarthy S, Cavender-Bares J. **2010**. Phylogenetic community structure in Minnesota oak savanna is influenced by spatial extent and environmental variation. *Ecography* **33**: 565–577.
Wilson KB, Baldocchi DD, Hanson PJ. **2000**. Spatial and seasonal variability of photosynthetic parameters and their relationship to leaf nitrogen in a deciduous forest. *Tree Physiology* **20**: 565–578.
Wirth C, Lichstein JW. **2009**. The Imprint of Species Turnover on Old-Growth Forest Carbon Balances - Insights From a Trait-Based Model of Forest Dynamics. In: Wirth C,, In: Gleixner G,, In: Heimann M, eds. Old-Growth Forests: Function, Fate and Value. Berlin, Heidelberg: Springer Berlin Heidelberg, 81–113.

Wright IJ, Ackerly DD, Bongers F, Harms KE, Ibarra-Manriquez G, Martinez-Ramos M, Mazer SJ, Muller-Landau HC, Paz H, Pitman NCA, *et al.* 2007. Relationships Among Ecologically Important Dimensions of Plant Trait Variation in Seven Neotropical Forests. *Annals of Botany* **99**: 1003–1015.
Wright SJ, Kitajima K, Kraft NJB, Reich PB, Wright IJ, Bunker DE, Condit R, Dalling JW, Davies SJ, Díaz S, *et al.* 2010. Functional traits and the growth–mortality trade-off in tropical trees. *Ecology* **91**: 3664–3674.
Wright IJ, Reich PB, Atkin OK, Lusk CH, Tjoelker MG, Westoby M. **2006**. Irradiance, temperature and rainfall influence leaf dark respiration in woody plants: evidence from comparisons across 20 sites. *New Phytologist* **169**: 309–319.
Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen JHC, Diemer M, *et al.* 2004. The worldwide leaf economics spectrum. *Nature* **428**: 821–827.
Wright JP, Sutton-Grier A. **2012**. Does the leaf economic spectrum hold within local species pools across varying environmental conditions? *Functional Ecology* **26**: 1390–1398.
Xu L, Baldocchi DD. **2003**. Seasonal trends in photosynthetic parameters and stomatal conductance of blue oak (Quercus douglasii) under prolonged summer drought and high temperature. *Tree Physiology* **23**: 865–877.
Yguel B, Bailey R, Tosh ND, Vialatte A, Vasseur C, Vitrac X, Jean F, Prinzing A. **2011**. Phytophagy on phylogenetically isolated trees: why hosts should escape their relatives. *Ecology Letters* **14**: 1117–1124.