

Quantifying the importance of local niche-based and stochastic processes to tropical tree community assembly

BILL SHIPLEY,^{1,4} C. E. TIMOTHY PAINE,² AND CHRISTOPHER BARALOTO³

¹Département de biologie, Université de Sherbrooke, Sherbrooke, Québec J1K 2R1 Canada

²Institute for Evolutionary and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

³INRA, UMR Écologie des Forêts de Guyane, 97387 Kourou Cedex, French Guiana

Abstract. Although niche-based and stochastic processes, including dispersal limitation and demographic stochasticity, can each contribute to community assembly, it is difficult to quantify the relative importance of each process in natural vegetation. Here, we extend Shipley's maxent model (Community Assembly by Trait Selection, CATS) for the prediction of relative abundances to incorporate both trait-based filtering and dispersal limitation from the larger landscape and develop a statistical decomposition of the proportions of the total information content of relative abundances in local communities that are attributable to trait-based filtering, dispersal limitation, and demographic stochasticity. We apply the method to tree communities in a mature, species-rich, tropical forest in French Guiana at 1-, 0.25- and 0.04-ha scales. Trait data consisted of species' means of 17 functional traits measured over both the entire meta-community and separately in each of nine 1-ha plots. Trait means calculated separately for each site always gave better predictions. There was clear evidence of trait-based filtering at all spatial scales. Trait-based filtering was the most important process at the 1-ha scale (34%), whereas demographic stochasticity was the most important at smaller scales (37–53%). Dispersal limitation from the meta-community was less important and approximately constant across scales (~9%), and there was also an unresolved association between site-specific traits and meta-community relative abundances. Our method allows one to quantify the relative importance of local niche-based and meta-community processes and demographic stochasticity during community assembly across spatial and temporal scales.

Key words: Community Assembly by Trait Selection, CATS; demographic stochasticity; dispersal limitation; environmental filtering; French Guiana; functional traits; maxent; neutral assembly; tropical forests.

INTRODUCTION

Neutral models (Bell 2000, Hubbell 2001) assume equal per capita demographic properties of all species. They thus presume that community assembly does not include niche-based processes, but arises via dispersal limitation from the meta-community and local demographic stochasticity. Deterministic models of community assembly (Lotka 1925, Volterra 1926, 1931, MacArthur 1972, Tilman 1982, 1988) represent the other extreme, in which niche-based factors dominate. Between these extremes lies a continuum of possibilities in which community assembly is influenced to varying degrees by local niche-based, stochastic, and landscape processes (Gravel et al. 2006). It has long been recognized (Tansley 1920) that the nonrandom processes of competition and abiotic environmental filtering interact with the unpredictable processes of seed arrival and demographic stochasticity to jointly determine community assembly. One reason why the debate

continues about which process dominates and under what conditions is that the relative importance of these contrasting processes has rarely (Cottenie 2005, Paine and Harms 2009) been quantified in a way that allows comparisons among studies (Condit et al. 2002, Clark and McLachlan 2003, Volkov et al. 2003, Gilbert and Lechowicz 2004, Harpole and Tilman 2006). To advance the debate we need a method of estimating the relative importance of these different processes that is comparable across locations, environments and spatial/temporal scales (Weiher et al. 2011). We developed such a method in this paper by extending the maxent model of trait-based community assembly presented in Shipley (2010b).

The model involves two spatial scales: the landscape containing a meta-community, whose species define the species pool, and one or more local communities within this landscape. The meta-community is thus composed of local communities that can potentially exchange propagules (Leibold et al. 2004), and a species is a member of the species pool if it has a nonzero probability of dispersing from the meta-community into a local community. However, a species will still be missing from a local community if dispersal is a

Manuscript received 24 May 2011; revised 1 November 2011; accepted 4 November 2011. Corresponding Editor: J. Cavender-Bares.

⁴ E-mail: Bill.Shipley@USherbrooke.ca

sufficiently rare event such that no immigrants have yet arrived, if the species is so poorly adapted to the local biotic and abiotic conditions that immigrants rarely reproduce and quickly die before being observed, or if demographic stochasticity results in local extinction.

Alternative model assumptions

First assume “local neutral” assembly, in which every species in the species pool has the same per capita probabilities of dispersing from the meta-community to the local community and of surviving, growing, and reproducing in the local community once present (Hubbell 2001). Given this assumption, species having more individuals in the meta-community will disperse more immigrants to the local community, even though the per capita probabilities of immigration are equal across species, and local dynamics would then proceed without regard to trait values or niche partitioning. This leads to the “neutral” prior expectation: that the expected relative abundance of each species in the local community is equal to its relative abundance in the meta-community. The actual relative abundance of each species in each local community will be a random walk around these expected values because of demographic stochasticity. As the spatial scale of the local community decreases, so also will the total number of individuals of various species in this local community. These smaller population sizes will increase the stochastic variation around the expected relative abundances of these species, and this will reduce the observed correlation between local relative abundance and meta-community relative abundance. In “local neutral” assembly, since trait values are unrelated to local fitness differences, they will also be independent of local relative abundances; therefore, the trait values of the average individual in the local community will be approximately equal to the average trait value in the meta-community. This is a restricted version of Hubbell’s (2001) neutral assumption because we make no assumptions about the causes of relative abundance at the landscape level, whereas Hubbell’s model assumes that neutral processes also determine the meta-community structure. In our “local neutral” model, the causes of relative abundance at the landscape level might be due to unanalyzed historical, phylogenetic, or biogeographic factors, but could also be due to niche-based processes. Thus, although dispersal limitation generates the similarity between the meta-community and local relative abundances in this “local neutral” model, it does not necessarily follow that dispersal limitation is independent of trait differences, only that trait differences among species co-occurring in the local community do not affect their local relative abundances. In other words, species’ traits provide no new information beyond that already encoded in the meta-community relative abundances.

Next, assume “local niche-based” assembly, in which the per capita probabilities of dispersal, survival, growth, and reproduction of a species in a local

community are completely determined by its functional traits in interaction with the local biotic and abiotic environment. If this were the case, then the relative abundance of the species in a local community would be entirely determined by the selective advantage of its traits, given the local environmental conditions. This follows because each species in the species pool has a nonzero probability of dispersing to the local community and subsequent population dynamics of the species, once it has arrived, would be determined by trait-based demographic properties of growth, survival, and reproduction. Once one knows the trait values of a species, then knowledge of its relative abundance in the larger meta-community would be irrelevant. This does not preclude demographic stochasticity from affecting the abundance of species with small population sizes. Those species having adaptively advantageous traits would have more individuals, and so the trait value of the average individual in the local community will be biased toward the more locally well-adapted species.

Finally, assume “hybrid” local assembly: The abundance of a species in the meta-community affects its probability of arrival (i.e., dispersal limitation) in a local community because more abundant species in the meta-community will tend to produce more propagules; therefore, meta-community abundances will contain relevant information. However, per capita probabilities of immigration will not be equal across species, and these differences will be associated with trait adaptations relevant to dispersal. Subsequent local dynamics are also shaped by trait values and local environmental conditions. In this case, knowledge of both the meta-community relative abundance and the trait values of a species would be informative in predicting its relative abundance in the local community. Species that are abundant in the meta-community could have somewhat higher abundances in the local community than their traits would predict due to mass flow (Shmida and Wilson 1985) or because a competitively superior species may be absent from the local community due to dispersal limitation. Species that are rare in the meta-community could have somewhat lower abundances in the local community than their traits would predict, because demographic stochasticity arising from a very small immigrant pool could force them to local extinction before their population sizes become large enough that trait selection dominates. The relative strength of dispersal limitation, trait-based filtering, and demographic stochasticity will determine the actual local community structure.

In this paper, we evaluate the relative importance of dispersal limitation, trait-based filtering, and demographic stochasticity in the assembly of species-rich tree communities of mature tropical rain forests in French Guiana for which extensive information is available on 17 functional traits. We further examine the extent to which the relative importance of each process is scale

TABLE 1. Summary of data used to fit the various models, by spatial scale: the number of separate local communities fit by the models and the median (minimum–maximum in parentheses) total number of individuals and trees species per community.

Level	Number of local communities	Median number of individuals	Median number of species
Meta-community	1	4557	657
1-ha scale	9	507 (406–632)	150 (145–202)
0.25-ha scale	36	123 (64–160)	69 (42–94)
0.04-ha scale	224	29 (5–32)	17 (4–28)

dependent by conducting our analyses at three spatial scales.

METHODS

The data (Baraloto et al. 2010a) come from a set of nine 1-ha tropical forest plots in French Guiana (see Plate 1) separated, on average, by 117 km (range: 0.5–300 km). Every tree of at least 10 cm in diameter at breast height was mapped and identified to species in 2007 and 2008. Because each individual was spatially referenced, we could also determine species relative abundances in subplots of 0.25 and 0.04 ha within each plot. These abundance data (number of individuals) were used to estimate relative abundances of each species both at the meta-community level (i.e., combining all nine plots) and at each of the three smaller spatial scales. Table 1 summarizes these data. There were 657 species in the meta-community, but only a small proportion of these were present in any single local community. Approximately 1500 tree species have been collected in all of French Guiana, including central and southern regions, which have a more distinctive flora (D. Sabatier and J.-F. Molino, unpublished data), so our sample likely represents at least 50% of the regional pool. The percentage of these 657 species that were missing from a single local community at the three spatial scales was 77% (1 ha), 89% (0.25 ha), and 97% (0.04 ha).

We measured values of 17 traits (Table 2) on almost every individual tree and calculated the average value

per trait for each of the 657 species in the species pool, as detailed in Baraloto et al. (2010a). This was done in two ways. First, we calculated the mean values per species averaged over all measured individuals per species over all nine plots (“meta-community trait means”). Second, we calculated the mean values per species averaged over all measured individuals per species in each 1-ha plot (“plot-specific trait means”); when a species was absent from a given plot its meta-community trait means were used. Although leaf traits were measured on every individual, wood traits and chemistry were not. We therefore estimated unobserved trait values using Multiple Imputation with Chained Equations (MICE), as implemented in the mice package of R (R Development Core Team 2008). Unobserved values were estimated through predictive mean matching using all other data as predictors. Community-weighted trait means, $\bar{t}_{jk} = \sum_{i=1}^S r_{ik} t_{ij}$, were calculated for each of the j traits in each of the k local communities (i.e., plots or subplots) based on each of the i species in the species pool; r_{ik} is the observed relative abundance of species i in local community k ; for plot-level trait means we used t_{ijk} , i.e., trait j of species i in plot k instead of t_{ij} . These average trait values were used as constraints in the maxent model.

Predicted relative abundances from the maxent model were obtained using the maxent function of R’s FD library (Laliberté and Shipley 2009). These predicted relative abundances are simultaneously maximum en-

TABLE 2. Functional traits measured on each of 657 tree species, along with units, number of individuals measured for traits, and median and range of observed species-mean trait values.

Functional trait	Unit	N	Median	Range in data set
Foliar $\delta^{13}\text{C}$ composition	‰	2946	−32.07	−36.44 to −25.67
Foliar C:N ratio	g/g	2947	24.01	7.79–59.77
Foliar K concentration	%	933	0.05	0.114–2.235
Foliar N concentration	%	2948	2.0	0.762–6.190
Foliar P concentration	%	933	0.1	0.024–0.251
Laminar chlorophyll content	$\mu\text{g}/\text{mm}^2$	4557	70.5	10.3–255.1
Laminar total surface area	cm^2	4557	72.1	2.032–643.70
Laminar toughness	N	4557	1.65	0.22–13.06
Leaf tissue density	g/cm^3	4540	0.042	0.008–0.287
Leaflet surface area	cm^2	4557	52.23	0.018–3218
NH_4^+ utilization	%	4557	47	1–100
Specific leaf area (SLA)	cm^2/g	4557	10.68	1.77–47.41
Trunk bark thickness	mm	3805	4	0.0–53.0
Trunk xylem density	g/cm^3	2844	0.65	0.23–0.98
Trunk xylem moisture content	%	2256	61	17–287
Twig bark thickness	mm	2369	2.0	0.07–7.31
Twig xylem density	g/cm^3	2390	0.62	0.19–0.96

trophy estimates and maximum likelihood estimates; this is explained more fully in the electronic appendix. Null probabilities for the statistical tests were obtained using the permutation test of Shipley (2010c) as implemented by the `maxent.test` function of the `FD` library. Because of the numerically intensive nature of this test when applied to such a large data set, we fixed a significance level of 0.05 and stopped the permutation runs when the 95% confidence intervals of the estimated null probabilities were outside the 0.05 level (Besag and Clifford 1991). For ease of comparison, and because the subplots smaller than 1 ha were clearly spatially dependent, we calculated the null probabilities separately for each of the nine plots even when the models were fit separately for each subplot, as described in Shipley (2010c). We used a maximally uninformative (uniform) prior when testing the pure local trait-based model and used a local neutral prior, consisting of the observed relative abundances at the meta-community level, for the hybrid model. These priors are explained in the following section.

Decomposing information content of a relative abundance distribution to infer contribution of traits, dispersal limitation, and other factors

We estimate the relative importance of stochastic and niche-based processes by modeling the pure local neutral, pure local trait-based, and the hybrid assumptions of community assembly using the maximum entropy CATS model (Community Assembly through Trait Selection) introduced in Shipley et al. (2006) and developed in detail in Shipley (2010b). Our approach to partitioning the total information content into components implied by these three models is similar to Borcard et al.'s (1992) method of partitioning the variance in constrained ordinations of abundance data into purely spatial, purely environmental, jointly spatial and environmental, and undetermined components. The Supplement summarizes the model.

The CATS model has three inputs: the trait matrix, the prior distribution, and the community-weighted trait means (trait constraints). The trait matrix lists the trait values of each species in the species pool; these values are species' means estimated over either the full meta-community or separately in each local community. The prior distribution is either the neutral prior (if non-trait-based dispersal limitation is assumed to occur) or a maximally uninformative uniform prior (if non-trait-based dispersal limitation is assumed not to occur). The trait constraints for a given local community are a vector of observed community-weighted trait means, which give the trait values of an average individual in the local community. The model finds the predicted relative abundances of each species in the species pool in each local community such that these predictions are consistent with the community-weighted trait means in the local community while simultaneously having

maximum relative entropy with respect to (thus as close as possible to) the given prior.

As such, the CATS model contrasts two opposing forces and allows one to estimate the relative importance of each. As the importance of local trait-based filtering increases, the proportion of the information contained in observed relative abundances that is accounted for by the trait model ("predictive ability") increases relative to the predictive ability given only the prior (either a uniform or neutral prior). As the importance of dispersal limitation from the landscape increases, the predictive ability of the model given only the neutral prior increases. Given a pure niche-based process of community assembly, the model will have maximal predictive ability (assuming that all relevant traits are incorporated), whereas the neutral prior alone will have no predictive ability. A strictly local neutral process of community assembly results in a model with no predictive ability beyond that provided solely by the neutral prior; in this scenario the predictive ability of the neutral prior will decrease as the importance of demographic stochasticity increases since the latter increases random variation around the expected neutral relative abundance distribution. Thus, the predictive ability of the model can be used to estimate the relative importance of trait-based and stochastic processes occurring during community assembly, and this predictive ability can be compared to alternative null distributions that incorporate different assumptions about the importance of dispersal limitation.

The proportion of the uncertainty in observed relative abundances that is accounted for by the model is quantified by Cameron and Windmeijer's (1997) generalization (R_{KL}^2 , Eq. 1), based on the Kullback-Leibler divergence, of the classic R^2 measure of goodness of fit between observed and model predicted values that is applicable in linear regression models with a normal error structure. This R_{KL}^2 generalization has all of the same properties possessed by the classic R^2 , and so R_{KL}^2 quantifies the proportion of uncertainty explained by the type of fitted exponential regression model that is implied by the CATS maximum entropy model. The supplement justifies this measure in detail. The total species pool is S (here, 657), the total number of local communities is C , whose value depends on the spatial scale, o_{ij} is the observed relative abundance of species i in local community j , p_{ij} is the model prediction of the relative abundance given the different assumptions of community assembly, and $q_{i,0} = 1/S$ is the predicted relative abundance given only the maximally uninformative prior:

$$R_{\text{KL}}^2 = 1 - \frac{\sum_{j=1}^C \sum_{i=1}^S o_{ij} \ln \left(\frac{o_{ij}}{p_{ij}} \right)}{\sum_{j=1}^C \sum_{i=1}^S o_{ij} \ln \left(\frac{o_{ij}}{q_{i,0}} \right)}. \quad (1)$$

We decomposed the total explained uncertainty in the observed relative abundances into separate components using the estimates of R_{KL}^2 given four different models, while correcting for model bias (Eqs. 2–5). The supplement develops and justifies the equations for this decomposition; only the final equations are presented here.

The data are fit to four versions of the full CATS model. The first model, estimating only model bias (with n representing null) and giving $\hat{R}_{\text{KL},n}^2(\text{uniform})$, is fit using a maximally uninformative (i.e., uniform) prior and permuted trait values. This removes any association between traits and relative abundances and excludes any contribution due to dispersal limitation (because a uniform prior assumes equal contributions of propagules of all species in the species pool) while maintaining any contribution to R_{KL}^2 due to model structure. The second model, estimating only dispersal limitation due to processes occurring at the landscape level and giving $\hat{R}_{\text{KL},n}^2(\text{neutral})$, is fit using a neutral prior and randomly permuted trait values. This again removes any association between traits and relative abundances but includes any contribution due to non-trait-related dispersal limitation. The third model, estimating only trait filtering and giving $R_{\text{KL}}^2(\text{uniform, traits})$, is fit using the maximally uninformative prior and non-permuted trait values. This incorporates any association between traits and relative abundances, but excludes any contribution due to non-trait-related dispersal limitation. The final model, estimating both dispersal limitation and trait filtering, is fit using the neutral prior and non-permuted trait values and gives $R_{\text{KL}}^2(\text{neutral, traits})$.

Eqs. 2 to 5 decompose the total explained uncertainty into five components: (1) model bias ($\hat{R}_{\text{KL},n}^2(\text{uniform})$), (2) the association due uniquely to trait-based filtering, (3) the association due uniquely to meta-community processes (non-trait-based dispersal), (4) the association due jointly to meta-community processes and local trait-based filtering, and (5) the remaining unexplained variance that can be provisionally ascribed to demographic stochasticity.

The proportion of biologically relevant uncertainty explained uniquely by local trait constraints is as follows:

$$\frac{R_{\text{KL}}^2(\text{neutral, traits}) - \hat{R}_{\text{KL},n}^2(\text{neutral})}{1 - \hat{R}_{\text{KL},n}^2(\text{uniform})}. \quad (2)$$

The proportion of biologically relevant uncertainty explained uniquely by the neutral prior is shown by

$$\frac{R_{\text{KL}}^2(\text{neutral, traits}) - \hat{R}_{\text{KL},n}^2(\text{uniform, traits})}{1 - \hat{R}_{\text{KL},n}^2(\text{uniform})}. \quad (3)$$

The proportion of biologically relevant uncertainty explained jointly by both the trait constraints and the neutral prior is as follows:

$$\begin{aligned} & [\hat{R}_{\text{KL},n}^2(\text{neutral}) + R_{\text{KL}}^2(\text{uniform, traits}) - \hat{R}_{\text{KL}}^2(\text{neutral, traits}) \\ & - \hat{R}_{\text{KL},n}^2(\text{uniform})] \div [1 - \hat{R}_{\text{KL},n}^2(\text{uniform})]. \end{aligned} \quad (4)$$

Biologically relevant unexplained information is shown by

$$\begin{aligned} & [1 - (R_{\text{KL}}^2(\text{neutral, traits}) - \hat{R}_{\text{KL},n}^2(\text{uniform})) \\ & - \hat{R}_{\text{KL},n}^2(\text{uniform})] \div [1 - \hat{R}_{\text{KL},n}^2(\text{uniform})] \\ & = \frac{1 - R_{\text{KL}}^2(\text{neutral, traits})}{1 - \hat{R}_{\text{KL},n}^2(\text{uniform})}. \end{aligned} \quad (5)$$

RESULTS

Trait-based filtering explained 72% of the information contained in relative abundances at the 1-ha scale, and this percentage decreased to 49% at the 0.04-ha scale (Table 3). The hybrid model, which takes both dispersal limitation and functional traits into account, explained 82% of the information in relative abundances at the 1-ha scale, decreasing to 55% at the 0.04-ha scale. The predictive ability of both the trait-only and the hybrid models was always greater using plot-specific trait means than meta-community trait means. Dispersal limitation, as inferred from the neutral prior, explained at least 39% of the information at all scales, but decreased slightly in importance as the spatial scale decreased.

Ecological inferences from the models are facilitated by considering only biologically relevant information on relative abundance. There was an association between the meta-community relative abundances (the neutral prior) and the site-specific trait constraints such that, after removing model bias, only 34% of the information at the 1-ha scale was uniquely attributable to niche-based processes. A similar amount of information (36%) was due jointly to traits and the neutral prior, and the two cannot be separated (Fig. 1). Also because of the meta-community abundance–local trait constraint association, the portion of the information explained uniquely by the neutral prior was only 10% at the 1-ha scale and remained approximately constant at all spatial scales. The portion of the total biologically relevant information that remained unexplained, and was tentatively ascribed to demographic stochasticity, was only 20% at the 1-ha scale, but increased to 50% at the 0.04-ha scale (Fig. 1).

DISCUSSION

The method developed here estimates the relative importance of local trait-based filtering, dispersal limitation, and demographic stochasticity to community assembly. The partitioning of information reflects the contribution of each process in determining the relative abundance of each species in the species pool within a

TABLE 3. Estimated proportions (R_{KL}^2) of the total information in the local observed relative abundances in 657 tree species in nine local sites in French Guiana that are accounted for by various maxent models.

R_{KL}^2 value	Scale		
	1 ha	0.25 ha	0.04 ha
Site-level trait means			
$\hat{R}_{KL,n}^2$ (uniform); model bias	0.066	0.081	0.162
$\hat{R}_{KL,n}^2$ (neutral); pure neutral	0.497	0.410	0.385
R_{KL}^2 (uniform, traits); pure trait selection	0.721 (9/9)	0.580 (9/9)	0.487 (9/9)
R_{KL}^2 (neutral, traits); hybrid model	0.815 (9/9)	0.658 (9/9)	0.554 (9/9)
Meta-community trait means			
$\hat{R}_{KL,n}^2$ (uniform); model bias	0.056	0.073	0.151
$\hat{R}_{KL,n}^2$ (neutral); pure neutral	0.501	0.406	0.387
R_{KL}^2 (uniform, traits); pure trait selection	0.146 (4/9)	0.134 (4/9)	0.179 (3/9)
R_{KL}^2 (neutral, traits); hybrid model	0.596 (7/9)	0.485 (7/9)	0.429 (6/9)

Notes: Shown are the results when the mean trait value of each species is calculated over all nine sites (meta-community trait means), and when the mean is calculated separately for each site (site-level trait means); a subscript n indicates “null.” Values in parentheses give the number of sites out of nine for which the trait constraints significantly improved the model over the prior alone.

local community. The minimal data requirements are measures of relative abundance in the local community and trait values of each species in the meta-community species pool. We applied this method to evaluate the degree to which the relative importance of these

processes changes with spatial scale. Our method would also facilitate evaluating the changes in relative importance of these processes along environmental gradients. Importantly, the relationship between these population processes and the resulting statistical model is explicit,

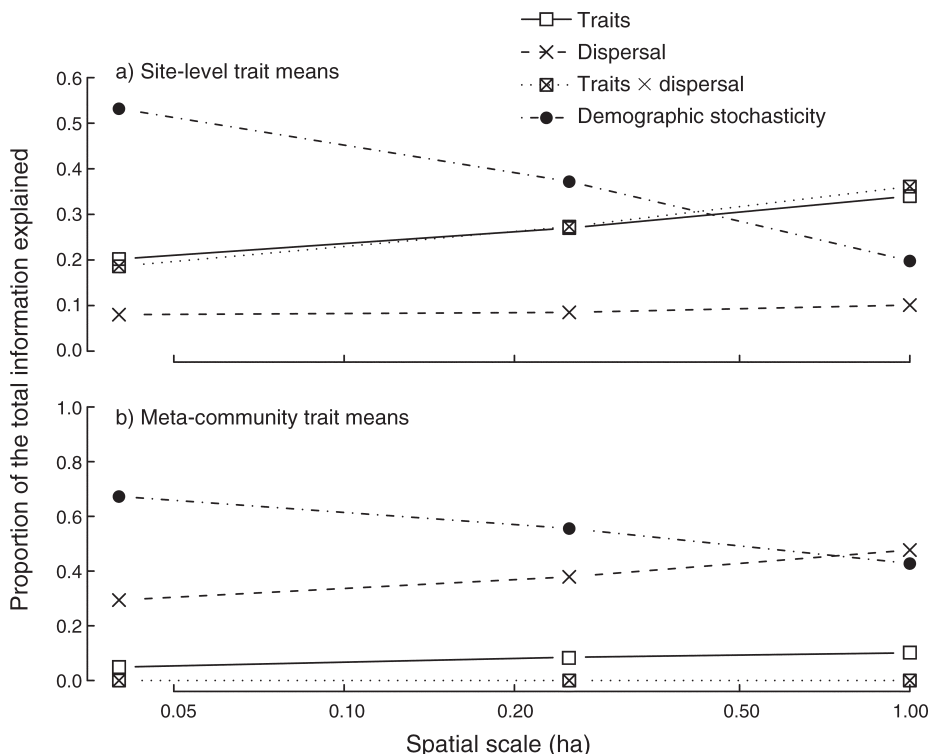


FIG. 1. Estimated proportion of the biologically relevant information in the local observed relative abundances of 657 tree species in nine local sites in French Guiana that is accounted for by various maxent models. The information is explained by local trait-based filtering, dispersal limitation from the meta-community, unexplained information attributed to demographic stochasticity and joint trait–dispersal effects. These are plotted at three spatial scales and based on species’ trait means calculated separately for each local community (“plot-specific trait means”) and trait means over the entire meta-community.

because the underlying model can be mathematically linked to population dynamics and quantitative selection on traits (Shipley 2010a, b:156–174).

Gilbert and Lechowicz (2004) and Cottenie (2005), based on Borcard et al. (1992), proposed a related method of decomposing species relative abundances that can also be linked to niche-based and neutral processes, but which does not use functional traits. This method requires information on (1) species abundance across many local communities (the community matrix **C**), (2) explicitly measured environmental variables in each local community (the environmental matrix **E**), and (3) the spatial configuration of the local communities (the spatial matrix **S**). Cottenie's (2005) variance decomposition consists of a (partial) redundancy analysis of **C** given either (1) **E**, (2) **S**, (3) **E + S**, (4) **E** controlling for **S** (**E|S**), or (5) **S** controlling for **E** (**S|E**). The assumption is that the partial redundancy analysis of **C** given the environmental matrix after controlling for the spatial matrix (**E|S**) reflects niche-based species filtering along the gradient (i.e., β diversity), while the partial redundancy analysis of **C** given the spatial matrix after controlling for the environmental matrix (**S|E**) reflects neutral processes or patch dynamics, but not niche-based processes. The **E + S** fraction, which is often the largest one, cannot be used to make conclusive statements about the relative importance of either **E** or **S**. However, unlike the CATS model upon which our method is based (Shipley 2010b:156–174), the statistical model has not been explicitly linked to these underlying population processes, and so one cannot know to what degree the assumptions are reasonable or robust. We could not apply Cottenie's (2005) method to our data because we could not identify clear abiotic gradients that differentiate our sites and that clearly drive the underlying dynamics. There are therefore three differences between our method and that proposed by Cottenie: (1) the type of data required, (2) the link between the statistical patterns and the demographic processes, and (3) the emphasis on explaining differences in relative abundance within a local community (as we do) vs. explaining differences between communities (Cottenie 2005). Cottenie's method and ours should therefore be viewed as partly complementary.

Our observational method contrasts with experimental approaches, including that of adding propagules of new species to a local community. The survival and reproduction of experimental colonists in a local community can provide evidence of dispersal limitation, but the method cannot quantify the relative importance of niche-based and neutral processes. By necessity, most such experimental studies have focused on taxa where dispersal is easily manipulated, where the organisms complete their life cycle rapidly, and where community dynamics are very rapid, as shown by the meta-analysis of Myers and Harms (2009). Although the link between pattern and process is more direct using this experimental approach, it is difficult to apply to many natural

systems and cannot provide a quantitative estimate of the relative importance of trait-based filtering and dispersal limitation. Furthermore, because such experiments are of relatively short duration, they can miss longer term processes of trait-based filtering (Paine and Harms 2009).

In previous applications of the CATS model to a diverse set of herbaceous plant communities, a strong signal of trait-based community assembly was detected, which suggested a dominant role for trait-based filtering (Shipley et al. 2006, Sonnier 2009, Mokany and Roxburgh 2010, Radovski 2010, Sonnier et al. 2010, Merow et al. 2011, Shipley et al. 2011). None of these studies, however, decomposed the information in species relative abundances and explicitly estimated the information explained by dispersal limitation. In the current study, contrastingly, trait-based filtering explained a maximum of only 34% of the information, and its contribution was less at smaller scales. The evidence of a predominance of trait-based processes provided by these previous studies is therefore still tentative, especially because such a sample of communities is much too small and unrepresentative to draw any general conclusions. In particular, community assembly in tropical forests seems to be particularly sensitive to stochastic processes; indeed, these are the communities that inspired Hubbell's (2001) neutral model. The large number of species in these forests that coexist at small spatial scales, without large differences in functional traits, and without pronounced environmental gradients, suggest that niche-based assembly processes will be weak. On the other hand, there is also empirical evidence of trait-based differences in performance and distribution among tropical tree species (Kraft et al. 2008, Poorter et al. 2008, Kraft and Ackerly 2010) that point to niche-based processes. What then is the relative importance of trait-based species' filtering vs. dispersal limitation vs. demographic stochasticity in structuring these tropical forests?

Our results help to answer this question. Local community structure in the studied tropical rain forests is consistent with neither a purely neutral nor a purely niche-based process. There was clear statistical evidence of trait-based filtering in the local communities, consistent with the earlier results of Paine et al. (2011), who studied the same forests. On the other hand, only 34% of the information in relative abundances was associated uniquely with trait-based filtering at the 1-ha scale, and this percentage decreased with decreasing spatial scale. Dispersal limitation was approximately constant across spatial scales, but was always less important than trait-based filtering based on the more informative plot-specific trait means. Because there was redundant information contained in the neutral prior and the site-level trait constraints, this proportion of the biologically relevant information cannot be unambiguously ascribed to only local niche-based or to dispersal limitation alone. However, the fact that the pure dispersal signal



PLATE 1. A fig tree, *Ficus* sp., in the lowland tropical rain forest at Nouragues Research Station, French Guiana. Photo credit: C. E. T. Paine.

remained essentially constant as the spatial scale of the local community increased, whereas the pure trait-based signal increased, suggests that the redundant information is due to species' traits affecting (or being correlated with) other landscape factors that affect relative abundances in the meta-community. It is impossible to identify these landscape factors given our data, but one possibility is that the relative abundance of the species in the meta-community partly reflects the relative abundance of the different environmental conditions in the landscape to which these species are adapted. This would explain both why there was redundant information between site-specific traits and meta-community abundances and why the pure dispersal limitation proportion of the information did not change much at the different spatial scales. However, it is also possible that the meta-community relative abundances simply reflect biogeographic processes of speciation (Hubbell 2001) and that our measured traits partly track this phylogenetic signal.

The unexplained variation, which we provisionally ascribe to demographic stochasticity, showed the opposite trend. At the 1-ha scale, with over 500 trees, on average, per local community, this unexplained variation was slightly less than that due only to dispersal limitation; but at the smallest spatial scale, with only 30 trees on average, it was three times more important. This is what one would expect if this unexplained variation was primarily due to demographic stochasticity: As population sizes decrease, chance fluctuations in individual birth and death rates would increasingly dominate population dynamics even if the species mean

probabilities of reproduction, survival, and growth were influenced by traits. This scale-dependence of the relative importance of neutral and niche processes has already been reported (Karst et al. 2005, Laliberté et al. 2009). However, the interpretation of the biologically unexplained information with demographic stochasticity must be made with care because if there are important unmeasured traits that were not strongly correlated with our measured traits then this would also contribute to the biologically unexplained information. Increasing the number of traits in the model will not artificially inflate the explained information because this source of bias is corrected using $\hat{R}_{KL,n}^2$ (uniform). Sonnier et al. (2011) describe how to select a parsimonious set of traits using the CATS model. Similarly, the importance of demographic stochasticity will be inflated if the values used in the neutral prior are poor estimates of the relative abundances of potential immigrants in the meta-community. Our estimate should therefore be considered as an upper bound on the importance of demographic stochasticity.

There are at least three sources of error that should be considered in our analysis. First, there might be strong but unknown environmental gradients that would result in different selection pressures on the traits within the same local community and that largely cancel out when averaged over the spatial scale of the entire local community. We have not found any such environmental gradients within the local communities and doubt that such strong gradients exist. Since the length of environmental gradients is often related to the spatial scale over which the environment is sampled (Bell and

Lechowicz 1994), then a local community defined at a larger spatial scale would be more likely to contain a larger environmental gradient within it. If so, then one would expect an increase in the strength of trait-based filtering at smaller spatial scales since any environmental gradients at such small scales would be less pronounced; this expectation is contrary to our results.

Second, we might have included many species in our species pool that have no possibility (as opposed to simply a low probability) of reaching our local communities. If the true species pool is greatly overestimated, then this necessarily leads to prediction errors in the maxent model (Sonnier et al. 2010). However, all but 30 species in our pool occur across the general region, all are likely capable of growing in all plots in the absence of competition, and there are no obvious nonbiological barriers that would make it impossible for any species to immigrate to each of the local communities. Properly identifying all those species that could reasonably disperse into, and survive, the abiotic conditions of the local communities (i.e., the species pool) is a difficult and unsolved problem. This is a weakness of the CATS model as currently implemented.

Finally, there might be strong trait-based filtering on traits not measured by us and that are largely independent of those that we did measure. The 17 measured functional traits expand on a well-accepted leaf economics spectrum (Wright et al. 2004) to define tropical tree tissue strategies (Baraloto et al. 2010b). Note that some of the previously cited studies using the CATS model detected much stronger associations between traits and relative abundance with as few as four or five traits. It is possible (even likely) that our analysis somewhat underestimates the degree of trait-based filtering; for instance, we do not have traits related to herbivore or pathogen defense or belowground traits, and these might be uncorrelated with our measured traits.

Despite the limitations discussed here, our method provides estimates of the relative importance of traits, dispersal limitation, and demographic stochasticity to the assembly of the tree communities in these species-rich forests. Our method should allow researchers to quantitatively compare across studies, locations, landscape features, and spatial/temporal scales. In this way we can move beyond the neutral/niche debate to study how such properties combine to determine the relative importance of niche-based and neutral processes during community assembly in different environmental, spatial, and temporal contexts.

ACKNOWLEDGMENTS

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Agence National de la Recherche (ANR) of France, the U.S. National Science Foundation (DEB 0743103), and an INRA Package grant to C. Baraloto. We thank all participants of the project BRIDGE who participated in field and laboratory

collection and treatment of specimens. Field research was facilitated by the Guyafor permanent plot network in French Guiana, which is managed by CIRAD, ONF, and CNRS. We thank Bruno Hérault, Eric Marcon, Vivien Rossi, Evan Weiber, and one anonymous reviewer for valuable comments.

LITERATURE CITED

- Baraloto, C., C. E. T. Paine, S. Patino, D. Bonal, B. Hérault, and J. Chave. 2010a. Functional trait variation and sampling strategies in species-rich plant communities. *Functional Ecology* 24:208–216.
- Baraloto, C., C. E. T. Paine, L. Poorter, J. Beauchene, D. Bonal, A.-M. Domenach, B. Hérault, S. Patiño, J.-C. Roggy, and J. Chave. 2010b. Decoupled leaf and stem economics in rain forest trees. *Ecology Letters* 13:1338–1347.
- Bell, G. 2000. The distribution of abundance in neutral communities. *American Naturalist* 155:606–617.
- Bell, G., and M. J. Lechowicz. 1994. The flip side: manifestations of how plants perceive patchiness at different scales. Pages 391–414 in M. M. Caldwell and R. W. Pearcy, editors. *Exploitation of environmental heterogeneity by plants. Ecophysiological processes above and below ground*. Academic Press, New York, New York, USA.
- Besag, J., and P. Clifford. 1991. Sequential Monte Carlo p-values. *Biometrika* 78:301–304.
- Borcard, D., P. Legendre, and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73:1045–1055.
- Cameron, C. A., and F. A. G. Windmeijer. 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* 77:329–342.
- Clark, J. S., and J. S. McLachlan. 2003. Stability of forest biodiversity. *Nature* 423:635–638.
- Condit, R., et al. 2002. Beta-diversity in tropical forest trees. *Science* 295:666–669.
- Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters* 8:1175–1182.
- Gilbert, B., and M. J. Lechowicz. 2004. Neutrality, niches, and dispersal in a temperate forest understory. *Proceedings of the National Academy of Sciences USA* 101:7651–7656.
- Gravel, D., D. Canham, M. Beaudet, and C. Messier. 2006. Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters* 9:399–409.
- Harpole, W. S., and D. Tilman. 2006. Non-neutral patterns of species abundance in grassland communities. *Ecology Letters* 9:15–23.
- Hubbell, S. P. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, New Jersey, USA.
- Karst, J., B. Gilbert, and M. J. Lechowicz. 2005. Fern community assembly: the roles of chance and the environment at local and intermediate scales. *Ecology* 86:2473–2486.
- Kraft, N. J. B., and D. D. Ackerly. 2010. Functional trait and phylogenetic tests of community assembly across spatial scales in an Amazonian forest. *Ecological Monographs* 80:401–422.
- Kraft, N. J. B., R. Valencia, and D. D. Ackerly. 2008. Functional traits and niche-based tree community assembly in an Amazonian forest. *Science* 322:580–582.
- Laliberté, E., A. Paquette, P. Legendre, and A. Bouchard. 2009. Assessing the scale-specific importance of niches and other spatial processes on beta diversity: A case study from a temperate forest. *Oecologia* 159:377–388.
- Laliberté, E., and B. Shipley. 2009. FD: Measuring functional diversity from multiple traits, and other tools for functional ecology. R package. R Foundation for Statistical Computing, Vienna, Austria.
- Leibold, M. A., M. Holyoak, N. Mouquet, P. Amarasekare, J. M. Chase, M. F. Hoopes, R. D. Holt, J. B. Shurin, R. Law, D. Tilman, M. Loreau, and A. Gonzalez. 2004. The

- metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters* 7:601–613.
- Lotka, A. J. 1925. *Elements of physical biology*. Williams and Williams, Baltimore, Maryland, USA.
- MacArthur, R. H. 1972. *Geographical ecology: patterns in the distribution of species*. Harper and Row, New York, New York, USA.
- Merow, C., A. Latimer, and John. J. Silander. 2011. Can entropy maximization use functional traits to explain species abundances? A comprehensive evaluation. *Ecology* 92:1523–1537.
- Mokany, K., and S. H. Roxburgh. 2010. The importance of spatial scale for trait-abundance relations. *Oikos* 119:1504–1514.
- Myers, J. A., and K. E. Harms. 2009. Seed arrival, ecological filters, and plant species richness: A meta-analysis. *Ecology Letters* 12:1250–1260.
- Paine, C. E. T., C. Baraloto, J. Chave, and B. Hérault. 2011. Functional traits of individual trees reveal ecological constraints on community assembly in tropical rain forests. *Oikos* 120:720–727.
- Paine, C. E. T., and K. E. Harms. 2009. Quantifying the effects of seed arrival and environmental conditions on tropical seedling community structure. *Oecologia* 160:139–150.
- Poorter, L., et al. 2008. Are functional traits good predictors of demographic rates? Evidence from five neotropical forests. *Ecology* 89:1908–1920.
- R Development Core Team. 2008. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Radovski, M. 2010. Testing a trait-based model of fern community assembly. McGill University, Montreal, Quebec, Canada.
- Shipley, B. 2010a. Community assembly, natural selection and maximum entropy models. *Oikos* 119:604–609.
- Shipley, B. 2010b. From plant traits to vegetation structure: chance and selection in the assembly of ecological communities. Cambridge University Press, Cambridge, UK.
- Shipley, B. 2010c. Inferential permutation tests for maximum entropy models in ecology. *Ecology* 91:2794–2805.
- Shipley, B., D. C. Laughlin, G. Sonnier, and R. Otfinowski. 2011. A strong test of the maximum entropy model of trait-based community assembly. *Ecology* 92:507–517.
- Shipley, B., D. Vile, and É. Garnier. 2006. From plant traits to plant communities: A statistical mechanistic approach to biodiversity. *Science* 314:812–814.
- Shmida, A., and M. V. Wilson. 1985. Biological determinants of species diversity. *Journal of Biogeography* 12:1–20.
- Sonnier, G. 2009. Des traits des espèces à la structure des communautés végétales, vers une approche fonctionnelle et prédictive de l'écologie des communautés. Université de Sherbrooke, Montreal, Quebec, Canada.
- Sonnier, G., B. Shipley, A. Fayolle, and M. L. Navas. 2011. Quantifying trait selection driving community assembly: A test in herbaceous plant communities under contrasted land use regimes. *Oikos* <http://dx.doi.org/10.1111/j.1600-0706.2011.19871.x>.
- Sonnier, G., B. Shipley, and M. L. Navas. 2010. Plant traits, species pools and the prediction of relative abundance in plant communities: a maximum entropy approach. *Journal of Vegetable Science* 21:318–331.
- Tansley, A. G. 1920. The classification of vegetation and the concept of development. *Journal of Ecology* 8:118–149.
- Tilman, D. 1982. *Resource competition and community structure*. Princeton University Press, Princeton, New Jersey, USA.
- Tilman, D. 1988. *Plant strategies and the dynamics and structure of plant communities*. Princeton University Press, Princeton, New Jersey, USA.
- Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2003. Neutral theory and relative species abundance in ecology. *Nature* 424:1035–1037.
- Volterra, V. 1926. Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Memorie di Ricerca Accademia dei Lincei, serie 6* 2:31–113.
- Volterra, V. 1931. *Leçons sur la théorie mathématique de la lutte pour la vie*. Gauthier-Villars, Paris, France.
- Weier, E., D. Freund, T. Bunton, A. Stefanski, T. Lee, and S. Bentivenga. 2011. Advances, challenges and a developing synthesis of ecological community assembly theory. *Philosophical Transactions of the Royal Society B* 366:2403–2413.
- Wright, I. J., et al. 2004. The worldwide leaf economics spectrum. *Nature* 428:821–827.

SUPPLEMENTAL MATERIAL

Appendix

The Community Assembly by Trait Selection (CATS) model (*Ecological Archives* E093-067-A1).

Supplement

R script and two R functions to implement the analyses described in our paper (*Ecological Archives* E093-067-S1).