Ecology, 92(7), 2011, pp. 1523–1537

Can entropy maximization use functional traits to explain species abundances? A comprehensive evaluation

Cory Merow,[1,3] Andrew M. Latimer,[2] and John A. Silander, Jr.[1]

[1]*Department of Ecology and Evolutionary Biology, University of Connecticut, 75 North Eagleville Road, Storrs, Connecticut 06269-3043 USA*
[2]*Department of Plant Sciences, University of California, 1 Shields Ave., Davis, California 95616 USA*

Abstract. Entropy maximization (EM) is a method that can link functional traits and community composition by predicting relative abundances of each species in a community using limited trait information. We developed a complementary suite of tests to examine the strengths and limitations of EM and the community-aggregated traits (CATs; i.e., weighted averages) on which it depends that can be applied to virtually any plant community data set. We show that suites of CATs can be used to differentiate communities and that EM can address the classic problem of characterizing ecological niches by quantifying constraints (CATs) on complex trait relationships in local communities. EM outperformed null models and comparable regression models in communities with different levels of dominance, diversity, and trait similarity. EM predicted well the abundance of the dominant species that drive community-level traits; it typically identified rarer species as such, although it struggled to predict the abundances of the rarest species in some cases. Predictions were sensitive to choice of traits, were substantially improved by using informative priors based on null models, and were robust to variation in trait measurement due to intraspecific variability or measurement error. We demonstrate how similarity in species' traits confounds predictions and provide guidelines for applying EM.

Key words: community-aggregated traits; community assembly; ecophysics; functional traits; fynbos; relative abundance; species abundance distributions.

Introduction

Research on functional traits and their relationship to abiotic gradients and to community composition and function has become a major focus of community ecology (e.g., Diaz et al. 1998, Ackerly and Cornwell 2007, Kraft et al. 2008). Due to the difficulty of parameterizing species-level models and to the appeal of finding general laws in community ecology, some ecologists have proposed working with functional traits, which are fewer in number and may encode information about both community and plant function (McGill et al. 2006). The entropy maximization model (EM) of Shipley et al. (2006) can potentially link such traits to community composition at the species level by predicting relative abundances of each species in a community using limited trait information. EM also offers a framework for quantifying the relationship of functional traits to environmental gradients (Shipley 2009*c*), for characterizing functional redundancy of species in communities, and for testing the sensitivity of trait-based inferences about community composition to trait variation from measurement error and intraspecific variation.

Manuscript received 10 June 2010; revised 14 October 2010; accepted 25 January 2011. Corresponding Editor: F. He.
[3] E-mail: cory.merow@uconn.edu

In this paper, we develop and apply a novel set of tests to critically evaluate the predictive power of EM under a variety of conditions. EM predicts the most likely community abundance patterns, given observed constraints, and goes a step beyond earlier models such as the neutral theory of biodiversity (Hubbell 2001) that seek to predict species abundance distributions while ignoring species identity (e.g., neutral theory does not determine which species is most abundant, only the abundance of the dominant species); EM predicts the relative species abundance distribution (RSA) of particular species based on their traits. The preliminary application of EM by Shipley et al. (2006) received early criticism (Marks and Muller-Landau 2007, Roxburgh and Mokany 2007, Haegeman and Loreau 2008, 2009), but despite recent optimism (He 2010, McGill and Nekola 2010) and further work by Shipley (2009*c*), it remains little tested. EM's potential accuracy is particularly appealing (Shipley et al. [2006] report $r^2 = 0.96$ between predicted and observed abundance) compared to other attempts to predict RSAs, which typically account for only 10–30% of the variance (McGill and Nekola 2010). However, many questions remain unanswered. How well does EM predict RSAs in communities with contrasting dominance and diversity patterns? How does it deal with species with similar traits? How do the quality and quantity of trait data affect predictions?

A number of distinct applications of entropy-based methods have emerged in recent ecological literature, and it is important to understand the context of our work in light of this new (to ecology) inductive framework. We distinguish between the maximum entropy formalism (MEF; Jaynes 1957, 2003), which is a general optimization method that can be applied to a variety of problems, and the specific application by Shipley et al. (2006) of MEF based on trait constraints in community ecology (denoted by EM). The common thread among MEF models is that the user minimizes bias in a predicted probability distribution with respect to a set of constraints on the moments of that distribution. One class of MEF models in ecology predicts macroecological patterns such as species abundance distributions or species–area relationships based on statistical considerations, and works in the currency of the number of individuals or species, or in total energy (Pueyo et al. 2007, Harte et al. 2008, 2009). Other applications predict the spatial distribution of a species based on the environmental conditions where it is observed ("Maxent"; Phillips et al. 2006). Other applications of MEF in ecology are also promising but have received relatively less attention and development (Dewar and Porté 2008, Volkov et al. 2009). The model that we consider predicts RSAs along environmental gradients based on species functional traits (Shipley et al. 2006, 2007, Haegeman and Loreau 2008, 2009, Shipley 2009*a*, *b*).

Conceptually, EM provides a model for habitat filtering (Keddy 1992, Díaz et al. 1998) that translates from plant functional traits to local community composition. The idea behind habitat filtering is that environmental conditions select for particular traits (community selection), and better-suited species pass through filters more readily, resulting in higher abundance than other species. Recent studies have shown strong correlations between convergent functional trait values and ecological strategies along environmental gradients (Westoby and Wright 2006, Ackerly and Cornwell 2007, McGill et al. 2007). EM connects models for functional trait variation along gradients back to species-level patterns by providing a mathematical formalism that uses the species' functional traits to predict their relative abundance.

The EM approach can be understood as follows. We sample a set of distinct local communities (e.g., across succession stages, elevational gradients, and so forth). Each local community type is sampled at multiple sites. We define the *local species pool* to contain all species observed at a collection of samples across a particular local community type. The *regional species pool* contains all species observed in all samples. We seek to predict the RSA in each local community from the regional pool, based on trait information for each species (e.g., leaf area, height). Species from the regional pool are indexed $i = 1, 2, ..., S$; traits are indexed $j = 1, 2, ..., T$; and for communities $k = 1, 2, \ldots, C$, we denote the RSA as vector $\mathbf{p}_k = (p_{1k}, p_{2k}, ..., p_{Sk})$. The information used to make the prediction is a set of T trait values measured for each of the S species, which define a trait vector $\mathbf{t}_i = (t_{i1}, t_{i2}, \ldots t_{iT})$ for each species i.

We summarize local community-level trait information as community-aggregated traits (CATs), which are the abundance-weighted average values of each trait j in each local community k (t_{jk}). In other words, we use these CATs to represent the operation of habitat filtering based on the typical trait values of plants found in each local community. For each local community k, the CATs can be calculated as follows:

$$\bar{t_{jk}} = \sum_{i=1}^{S} p_{ik} t_{ij} \qquad \text{for } j = 1, 2, \ldots T \qquad k = 1, 2, \ldots, C. \quad (1)$$

Because CATs have been shown to describe community-level variation along abiotic gradients at multiple scales, we can take advantage of the fact that they generally vary relatively smoothly along environmental gradients (Garnier et al. 2004, Ackerly and Cornwell 2007, Lavorel et al. 2008 [and 41 references in their Appendix 1]) to predict CATs. CATs can be predicted for each local community by using splines fit to empirical CATs calculated from (1) with multiple communities along the gradient. Using these imputed CATs, EM predicts the RSA in each community, $\mathbf{p}_k$.

EM requires the user to specify the possible system states, prior distribution for abundance, and constraints (Haegeman and Loreau 2009; see Appendix A). First, the system states are given by $\mathbf{p}_k$, the relative abundance of each species in the community. Second, the user must specify a *prior* distribution $\mathbf{q}_k$ (a vector for community k) that reflects the state of knowledge about the RSA before imposing constraints. Previous applications of EM have placed a uniform prior on $\mathbf{p}_k$; this implies that all species in the potential species pool are equally likely to occur in a local community. Third, for each trait, $\mathbf{p}_k$ must satisfy linear constraints based on the CATs in Eq. 1. In addition, the $\mathbf{p}_{ik}$ must sum (over i) to one and be nonnegative. MEF finds the vector $\mathbf{p}_k$ that maximizes the local community's relative entropy function (i.e., minimizing information and/or maximizing similarity to the prior):

$$-\sum_{i}^{S} p_{ik} \ln\left(\frac{p_{ik}}{q_{ik}}\right) \quad (2)$$

subject to the constraints (Eq. 1) and normalization. Given t_{jk}, the maximization is readily solved for $\mathbf{p}_k$ using the method of Lagrange multipliers, to yield the following (Jaynes 2003):

$$\hat{p}_{ik} = q_{ik} \exp\left(\lambda_{0k} - \sum_{j=1}^{T} \lambda_{jk} t_{ij}\right) \Bigg/ \sum_{i=1}^{S} \exp\left(\lambda_{0k} - \sum_{j=1}^{T} \lambda_{jk} t_{ij}\right) \quad (3)$$

where the λ_j are the Lagrange multipliers. There is one λ_j for each trait and λ_0 corresponds to the normalization constraint on $\mathbf{p}_k$. When the traits are standardized such

that each is restricted to the interval [0,1] (Petchey and Gaston 2006), the absolute values of the λ_j indicate the relative predictive strength of the functional traits in predicting species abundances (Phillips et al. 2006). This method is applied independently to each local community (Shipley et al. 2006).

EM can be conceptualized in terms of two components. First, the CATs describe the *constraint component*, which defines a group of RSAs that are consistent with functional constraints (Eq. 1) and are denoted as the *feasible set*. This constraint component restricts the possible combinations of trait vectors and can be conceptualized as the niche constraints on the assemblages of species that can occupy a particular local community (see *Discussion* for an examination of this assertion). Entropy is not involved in these constraints. Second, the *entropy component* selects the RSA with maximum similarity to the prior from the feasible set. The maximum entropy condition ensures that the RSA contains the minimum information about the species identity of a randomly selected unit of biomass from the local community. EM therefore provides the only unbiased RSA with respect to the prior and observed constraints (Shipley 2009*a*). If EM performs well, it can be either because the traits are extremely constraining, or because the community tends more toward the prior than would be required by trait constraints alone. Note, however, that the justification for maximizing entropy is based on information theory and has no interpretation in terms of ecological mechanisms.

We assessed the predictive value of functional traits to determine abundance using a suite of complementary tests. We applied the model to the same set of successional sequences of low-diversity, abandoned French vineyards that Shipley et al. (2006) studied ("vineyard data"), and also an elevational transect in the highly diverse fynbos shrubland of the Cape Floristic Region of South Africa (Proches et al. 2003). The fynbos data set potentially offers further insight into EM because it is more diverse, contains a wider array of dominance patterns, was collected over larger spatial scales, and contains more rare and functionally similar species than the vineyard data. We first test the predictability of CATs along a gradient. We then assess the power of EM to use traits to predict RSAs using null models, cross validation, and comparison to regression models. By comparing predictions made using regional vs. local species pools, and incorporating prior information reflecting propagule availability, we use EM to explore how recruitment limitation affects local community assembly. We also test the robustness of predictions to functional trait variability. Finally, we discuss how EM's predictions could shed light on functional redundancy and provide guidelines for applying EM to other data sets.

Methods

Data

We applied entropy maximization to two contrasting species abundance data sets. The first comes from 12 old fields (local community samples) in a Mediterranean climate region in southern France (Garnier et al. 2004), as analyzed by Shipley et al. (2006). These fields were formerly vineyards abandoned at different times and constitute a chronosequence (2, 2, 7, 8, 8, 11, 12, 26, 29, 35, 40, and 42 years since abandonment). These samples consist of single 0.5×0.5 m sites. Wiegart (1962) has shown that this is an appropriate, minimum sample area for such herbaceous communities. Following Shipley et al. (2006), we truncated the data to retain 30 species, based on abundance in the regional community, which constitute 80% of the regional biomass. The local community samples constituting the chronosequence consisted of 12, 9, 12, 10, 8, 15, 13, 4, 4, 4, 6, and 4 species, respectively. The following traits were measured for each species, defining a trait vector (supplying t_{ij} in Eq. 1): proportion perennial, seed number, seed maturation date, specific leaf area, aboveground vegetative mass, stem mass, leaf mass, and height (see Garnier et al. 2004).

We then examined eight fynbos communities sampled along an elevational gradient in the Baviaanskloof Mountains of the Cape Floristic Region of South Africa (Proches et al. 2003). This data set consisted of 42 5×10 m local community samples along a 50-km transect (Appendix B: Fig. B1), partitioned into eight local communities (Appendix B: Fig. B2) spanning elevational range intervals of 120 m each, from 250 m to 1205 m. In contrast to the treatment of Shipley et al. (2006; see *Discussion*), local communities were characterized by multiple samples within the same elevation range (5, 5, 7, 4, 4, 7, 7, and 3 samples per local community, by increasing elevation), with individual samples separated spatially (without replication) by hundreds to thousands of meters. The diversity and spatial heterogeneity in fynbos necessitates multiple samples of this size to accurately characterize communities. We retained the 43 most abundant species in the region, based on percent cover, constituting 77% of regional cover (in parallel to Shipley et al. 2006). The local communities are much more diverse than the vineyard communities, consisting of 18, 21, 28, 19, 28, 27, 25, and 21 species, respectively. There is a strong gradient in this system: temperature decreases and precipitation increases with elevation (Rebelo et al. 2006). The average daily minimum temperature of the coldest month (July) ranges from 1° to 6°C and mean annual precipitation ranges from 300 to 700 mm. Using Goldblatt and Manning's (2000) Conspectus of the Cape Floristic Region, we obtained the following traits to provide information on plant function perennial (yes/no), succulent (yes/no), plant height, leaf longevity (<1 year, 1–3 years, >3 years), flowering duration, and pubescence (yes/no, on either

TABLE 1. Methods used to explore entropy maximization (EM) and general conclusions.

Quantity varied	Questions	Main conclusions
CAT predictability	Are CATs predictable along environmental gradients?	CATs correlated strongly with gradients.
Predicting CATs	How sensitive are abundance predictions to deviations from observed CATs?	Predictions were very sensitive to CAT accuracy indicating their predictive value.
		Abundance predictions decreased under CAT cross validation but predicted CATs remained sufficient to differentiate communities.
Richness	How well does EM perform with varying richness?	Better predictions in simpler communities.
		Predictions improved as richness decreased, but decreased below a certain threshold when CATs could not be easily fulfilled by remaining species.
Trait similarity	How does trait vector similarity affect prediction accuracy?	Functional similarity within guilds reduces predictive power.
		Effects of trait similarity on EM depend on complex combinations of species.
Number of traits	How much trait information is necessary or sufficient?	Steady decline in prediction accuracy when fewer traits were used.
	Which traits provide the most predictive power?	Lagrange multipliers identified the most important traits.
	Do quantitative vs. categorical traits matter?	Categorical and quantitative traits have comparable predictive power.
Priors	How do different priors affect prediction accuracy?	Regional priors were better than uniform priors, indicating importance of regional abundance for local abundance.
	Can priors be used to model dispersal limitation effects?	Subregional priors performed best in Fynbos, suggesting importance of dispersal limitation in that system.
Trait accuracy	Are predictions sensitive to intraspecific trait variation or measurement error at the species level?	Predictions were robust to perturbations in trait accuracy at 5% and deviated more substantially with 20% perturbations.
Null models	What are the appropriate null models to compare against predictions?	Predictions were better than 98.4% of null models except in two cases.
	Does EM perform better than null models?	
EM vs. regression	Does EM perform better than regression models, using the same information?	EM outperformed regression models in all cases.

Note: CATs are community-aggregated traits (weighted average).

stems or leaves). We measured herbarium specimens to obtain leaf width, leaf perimeter^2/area, leaf area/twig basal diameter, and stem length/stem basal diameter.

Analysis

We will outline the questions about EM that our analyses address and the methods used to answer them, as summarized in Table 1. We begin by describing three tests we created that are relevant to a number of the questions. These tests are all performed using a uniform prior, for comparison to the results of Shipley et al. (2006) except where explicitly noted otherwise. All three tests (T) start with the observed CATs (see Eq. 1):

1) T_{loc} predicts RSAs for each local community given observed CATs, selecting only from species actually observed in the respective *local* community.

2) T_{reg} predicts RSAs for each local community given observed CATs, selecting from *all* species in the *regional* pool. Comparison to T_{loc} gives a measure of the species' similarity in trait space. T_{reg} is equivalent to the analysis in Shipley et al. (2006: Fig. 2).

3) T_{sm} uses the same species pool as T_{reg} to predict CATs from *smoothing* splines fit to the observed CATS for each local community along the gradient (instead of using observed CATs). This tests the finding of Shipley et al. (2006) that RSAs predictions are similarly accurate whether observed or predicted CATs are used. This test is analogous to the second analysis of Shipley et al. (2006: Fig. 3), although more robust because we do not employ smoothing techniques on the abundances.

Are CATs predictable along environmental gradients?—Shipley et al. (2006) demonstrated that CATs varied regularly for the vineyard data so we asked whether that predictability generalized to more diverse systems with different gradients. For both data sets, we predicted CATs along a gradient by fitting a smoothing cubic spline regression (Schumaker 2007) through the empirical values, and we reported the correlation of empirical values with regression predictions (Fig. 1). If CATs varied regularly along the gradient, we presumed that they described the effects of habitat filtering in response to the gradient (Petchey and Gaston 2006).

How sensitive are abundance predictions to deviations from observed CATs?—We compared the results of T_{sm}, which uses predicted CATs at all points along the gradient, to T_{reg}, which uses observed CATs to evaluate the sensitivity to CAT values (Fig. 2). We also used hold-one-out cross validation for CAT predictions

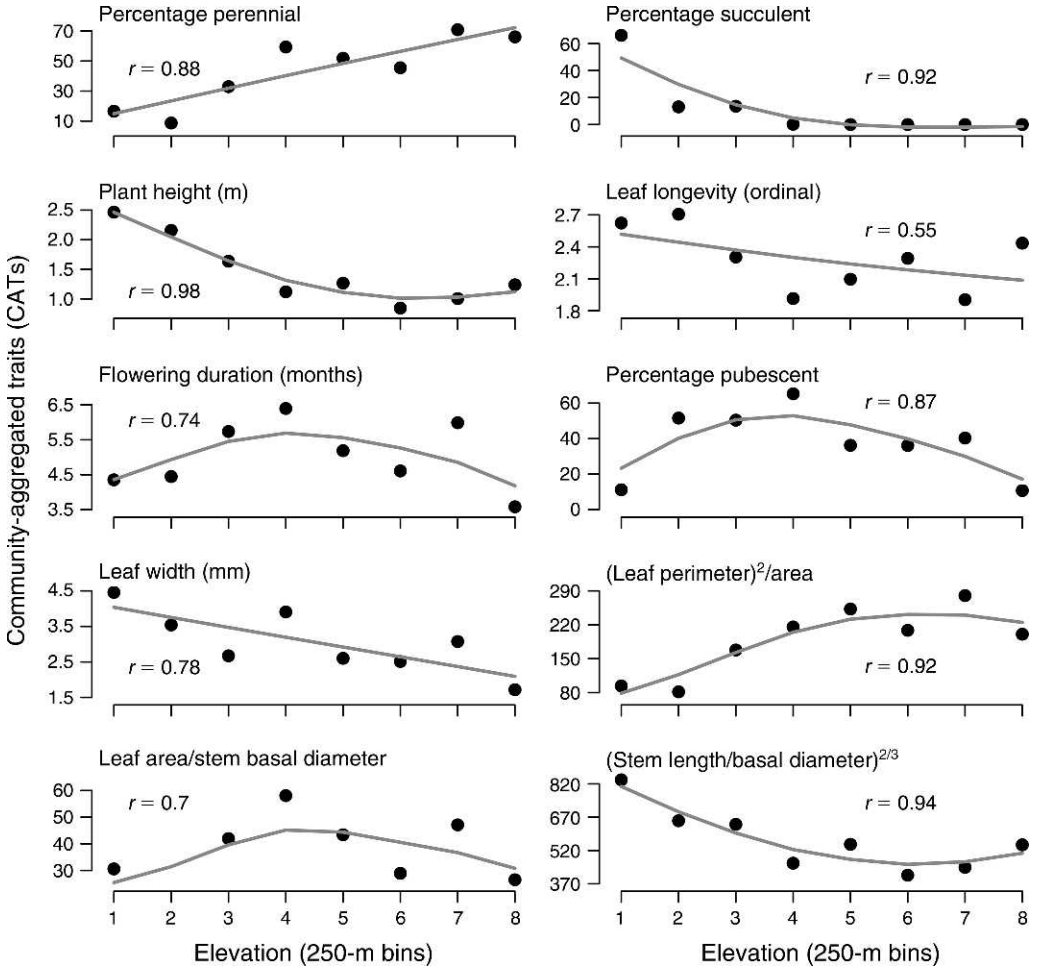

FIG. 1. Environmental heterogeneity as reflected in the variation of observed (points) and fitted (lines) CATs (community-aggregated traits) along the elevational gradient for the South African fynbos data. Elevation has been grouped into eight bins, each spanning 250 m. CATs are calculated as the abundance-weighted average of traits in the local community. Leaf and stem measurements were made in millimeters.

following Marks and Muller-Landau (2007). Each local community was sequentially withheld from the spline fit to CATs along the gradient and its CATs were predicted.

How well does EM perform with varying species richness?—We began by comparing the results for the vineyard and fynbos data, which exhibited contrasting patterns of dominance and diversity at the local community level (Appendix D: Figs. D8–11). To evaluate the effects of local vs. regional species pools, we compared the predictions of T_{loc} and T_{reg}. We also implemented T_{loc}, T_{reg}, and T_{sm} while varying the number of species used to define a local community (Fig. 3). The number of species in each data set was varied (vineyard, 9–30 species; fynbos, 11–43 species) by removing the rarest species in the truncated regional community at each step. We calculated prediction surfaces that show model fit on the *z*-axis, number of species on the *x*-axis, and number of traits on the *y*-axis in two ways: (1) using the CATs calculated from all species, and (2) recalculating CATs each time the species pool was reduced. Case 1 corresponds to using the most complete information possible, which is necessary when CATs are predicted from splines. Case 2 corresponds to predictions focused on dominant species, i.e., when sampling is a problem and some species are omitted from the model (as shown in Appendix D).

How does trait vector similarity affect prediction accuracy?—Assessing the effect of trait vector similarity on predictions is challenging because there are many possible linear combinations of trait vectors that are substitutable for a particular species' trait vector rather than simple pairs of similar species. One cannot isolate the effect of adding or subtracting one potentially similar species from the species pool because predictions depend on all other present species. Thus we summarized trait vector similarity based on (1) community-level and (2) species-level trait vector similarity. For community-level similarity, we compared the prediction accuracy under T_{loc} (measured by h) for each local

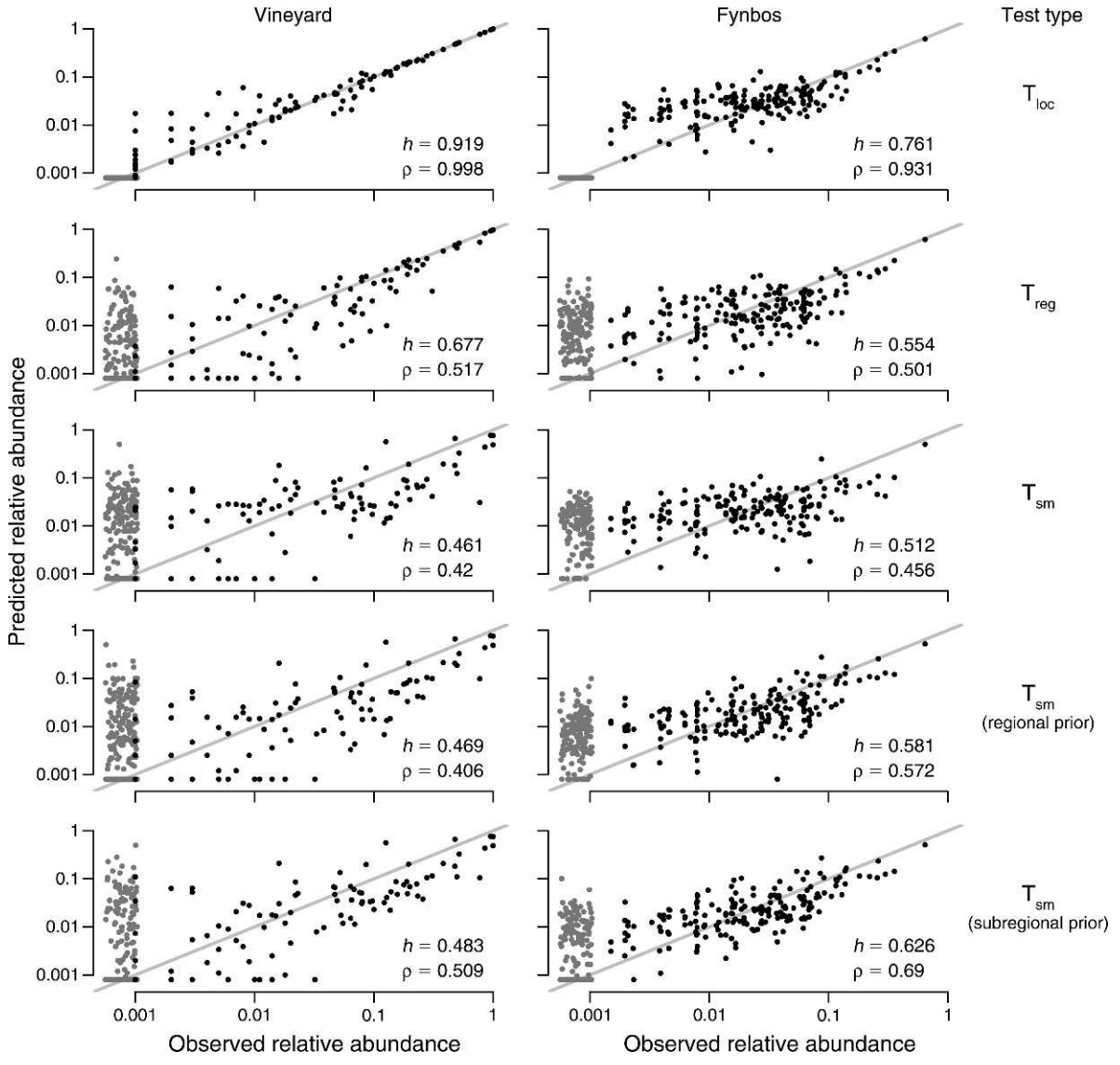

FIG. 2. Variation of French vineyard and South African fynbos prediction success (h, ρ) for different tests of entropy maximization, EM. Predictions for all local communities (8 in the vineyard and 12 in the fynbos) are shown on the same plot. Tests were applied with the regional species pool and all traits. T_{loc} (local) fits the model using only species observed in the local community while T_{reg} (regional) and T_{sm} (smoothed) use all species from the regional community. T_{sm} uses predicted CATs while T_{loc} and T_{reg} use observed CATs. Uniform priors are used for the first three tests shown; priors generated from the subregional and regional species pools are also used with T_{sm}. Gray dots have been jittered along the abscissa to show their density. Gray lines indicate 1:1 agreement between predicted and observed relative abundance of species (log–log scale); ρ is the Pearson correlation coefficient between ranks.

community (12 in the vineyard, 8 in the fynbos) to the trait vector dissimilarity in the respective local community. The trait vector dissimilarity was computed as the mean of all pairwise Manhattan distances of locally observed species divided by the number of traits (to make data sets comparable).

To examine species-level similarity, we tested the hypothesis that if nearest neighbor distance in trait space decreases, then abundance predictions become less accurate. We measured nearest neighbor distance with the Manhattan distance and prediction accuracy as the absolute error between predicted and observed abundance for each species. We then plotted the changes in the nearest neighbor distances and prediction error changed between T_{loc} and T_{reg}. The nearest neighbor distances typically decreased from T_{loc} to T_{reg} because T_{reg} contains more species.

How much trait information is necessary or sufficient? Which traits provide the most predictive power? Do quantitative vs. categorical traits matter?—To determine the predictive value of different numbers of traits, for each of T_{loc}, T_{reg}, and T_{sm}, the quantity of traits was

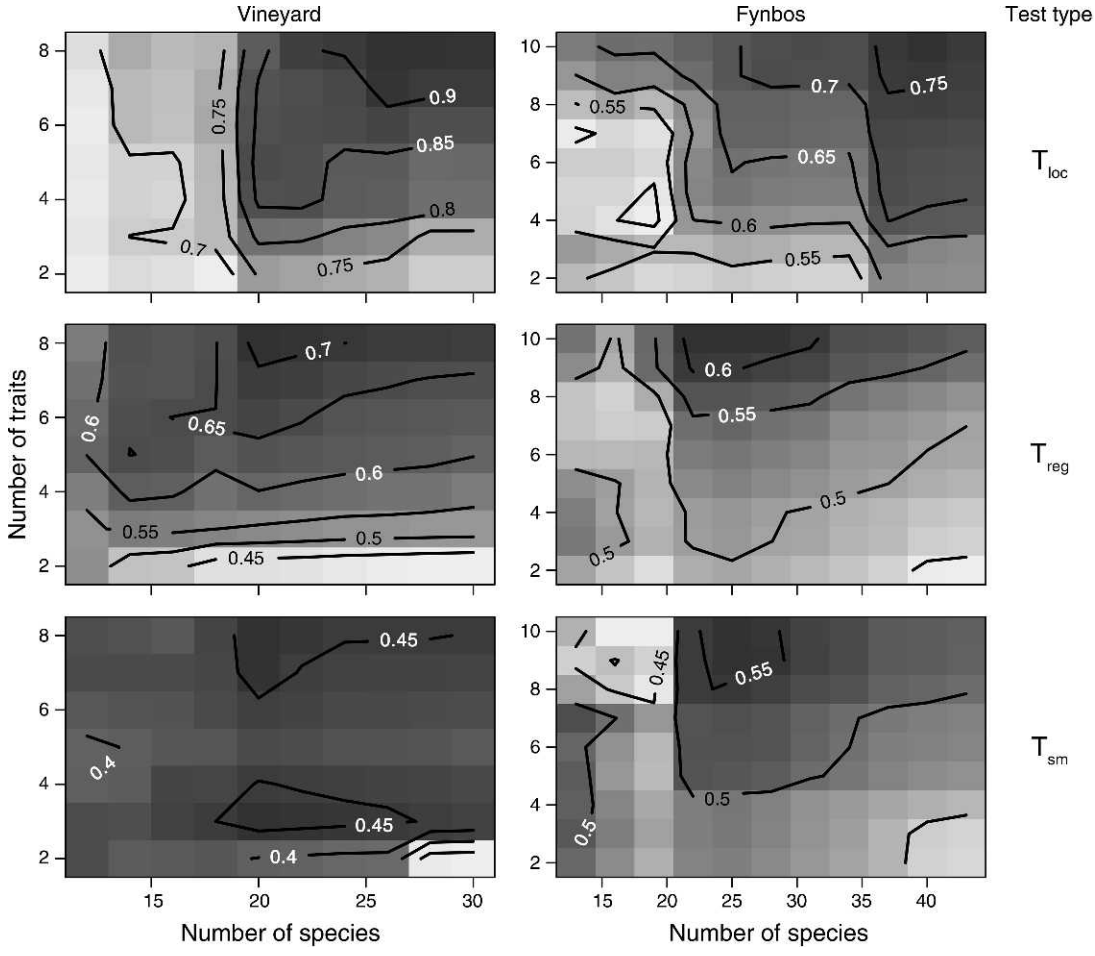

FIG. 3. Prediction success (h) for varying numbers of species and traits under three tests (T_{loc}, T_{reg}, T_{sm}). The trait with the lowest mean LaGrange multiplier across all local communities (when using all species) was omitted at each step. The species with the lowest regional abundance was omitted at each step. All three tests then proceeded as if the regional species pool were this truncated pool. Initially, h increases with decreasing number of species as trait vector similarity decreases. Then h decreases when very few species remain because important species are missing that are necessary to fulfill the CATs.

varied by removing those with the lowest predictive value, determined as the lowest mean Lagrange multiplier across all local communities. Lagrange multipliers indicate the predictive values of traits.

How do different priors affect prediction accuracy? Can priors be used to model dispersal limitation effects?—We considered the effects of using different priors to reflect different empirical expectations. No studies of which we are aware predict uniform RSAs; thus a better null model than equal a priori abundance should account for similarity to the surrounding regional community. In the absence of any expectations about niche relationships, a species might be abundant in a local community simply because it is abundant in the regional community. This could occur even if all seeds arrive randomly in the local community; more abundant species would provide more seeds. To account for this phenomenon, we used the regional RSA as the prior.

The prior can also reflect other mechanisms that might alter abundance expectations that are not included as traits, such as dispersal limitation. For example, if random seed arrival is expected to be important, as previously discussed, but at smaller spatial scales than the entire region, one could use the RSA composed from plots nearby the local community of interest. We did this by constructing prior RSAs from only the local communities adjacent to the community of interest along the environmental gradient; i.e., the prior RSA for four-year-old vineyards comes from vineyards between 0 and 8 years old. We termed such RSAs *subregional* RSAs and calculated them by averaging the RSAs of five local communities ranging from two below to two above the focal community along the gradient (e.g., for elevational community 4, the prior was the average RSA from communities 2–6). To account for sampling error, we generated random samples from the regional and subregional RSAs and used these as priors. We compared these priors to random samples from a uniform distribution to understand the sensitivity of predictions to different

priors. See Appendix D: Figs. D14–17 for an alternate way to construct priors based lognormal and exponential distributions.

Are predictions sensitive to intraspecific trait variation or measurement error at the species level?—To test the sensitivity of predictions to trait measurement, we used a perturbation analysis on the traits. Because we have no information on intraspecific variance or measurement error for the quantitative traits we used, we examined the consequences of changing the traits' values (perturbation) by both small (within 5% of observed values) and large amounts (within 20% of observed values) on the model predictions. Larger perturbations address the recent focus on large intraspecific variability in functional ecology; these perturbations cover the typical interquartile range of traits studied by Albert et al. (2010). Perturbed traits were chosen for each species from a uniform distribution ranging from 5 (or 20)% above to 5 (or 20)% below observed trait values. Next, we proceeded in one of two ways. First, we *recalculated* the CATs using the perturbed traits, and then applied EM. This examined whether the predictions were an idiosyncratic product of our data or only relevant for the small perturbations. Second, we used the perturbed traits with the *observed* CATs. This tested the sensitivity of the predictions to the intraspecific variability and measurement error. Note that we did not perturb categorical traits (1 in the vineyard, 4 in the fynbos) because it was unclear how to treat these types of traits and it seemed reasonable to assume that they were measured without error. We ran 500 model replicates for each of the three tests (T_{loc}, T_{reg}, and T_{sm}).

Does EM perform better than null models?—Predictions were compared to null models based on permutation of the trait vectors (Roxburgh and Mokany 2007, Shipley et al. 2007). Under the null hypothesis, traits do not influence species' abundance (Manly 2006, Shipley et al. 2007). Thus we randomly reassigned trait vectors to species, recalculated the CATs, and predicted abundances. The only attribute changed in the null model for a particular data set or test was the trait vector associated with each species. Null models were computed for each test shown in Fig. 2. Importantly, this null model retains the covariance among traits and maintains the same trait vector similarity as the observed data. The null hypothesis contrasts the expectation that if species' traits respond to the environment, then species with similar traits will likely co-occur. The alternative of using completely random traits (Roxburgh and Mokany 2007) or randomizing a single trait at a time is not appropriate when covariance is strong among traits because this allows for biologically unrealistic species and substantially reduces trait vector similarity. Such unrealistic trait vectors will be much more restricted by the constraint component and tend to overestimate the support for the null hypothesis. We report the proportion of predictions better than the true model obtained from 500 such permutations.

Does EM perform better than regression models, using the same information?—Even if EM makes good predictions, if it performs no better than standard regression models, there would be less reason to use it. However constructing an appropriate regression model with relative abundance as the response variable is not straightforward; predicted abundances must be constrained to the interval [0,1] and to sum to unity. The multinomial logistic regression suggested by He (2010) represents a viable (and arguably preferable) alternative to EM when count data are available, but this does not apply when only relative abundance data are available, as is often the case for plant species surveys.

We constructed linear regression models with the response variable given by the ratio of the relative abundance of each species to a reference species. Optimal Box-Cox (BC) transformations of the response were used to maximize the normality of the residuals. Violation of the normality assumption is possible for this model; more serious violation was apparent for T_{loc}, but for T_{reg} and T_{sm}, deviations from normality were minimal for both data sets. Constructing optimal regression models that are comparable to EM is an important area for future research (He 2010); our models represent a first step. Defining p_i as the relative abundance of the ith species and p^* as the reference species, we can write the model as

$$\mathrm{BC}(p_i/p^*) \sim \mathbf{X}\boldsymbol{\beta} + \varepsilon_i$$

where $\mathbf{X}$ is the trait matrix, $\boldsymbol{\beta}$ is the vector of regression coefficients, and the ε_i are the Gaussian distributed errors. We applied models to each local community separately (to make a direct comparison to EM) and chose the most locally abundant species as the reference species in each community. This formulation allowed us to scale relative abundances, based on the predicted ratios, to sum to unity. The optimal Box-Cox transformation for the fynbos data corresponded to a logarithmic transformation, so this required that we add a small quantity to relative abundance values that were observed to be zero. For relative abundance values of zero, we used 10^{-6} to circumvent the log(0) problem. Zero abundances are likely to reflect Preston's veil-line concept (Preston 1948, Chisholm 2007), i.e., species too rare to be sampled (smaller values led to the worst predictions). Note that one strength of EM is that it does not require such ad hoc assumptions. Finally, we chose two different sets of predictors: (1) raw traits (model R1), and (2) the squared difference between each species' traits and CATs for the each community (model R2). The latter provided a parallel to the constraints imposed by the CATs in EM, because a species' relationship to the average trait set is expected to determine its abundance. For both models, we made predictions based on T_{loc}, T_{reg}, and T_{sm} (R2 only).

Goodness of fit.—We assessed model results by plotting predicted vs. observed abundance. All local communities are shown on the same plot (Fig. 2);

Appendix D: Figs. D8–11 show predictions for each local community separately. We used two different metrics to compare predictions with observations that avoid the biases of the more traditional R^2 (cf. Kvalseth 1985, McGill 2003). R^2 values are sensitive to outliers (i.e., most abundant species), insensitive to relative error in points near the origin (i.e., rare species), and are calculated with respect to the least-square regression line, *not* the relevant 1:1 line. We report the information-theory-based Hellinger divergence (H) because it offers a general way to compare two probability distributions (Ali and Silvey 1966). The Hellinger distance measures the distance between two discrete distributions, the observed (O, with elements o_i) and predicted (P, with elements p_i) RSAs, as follows:

$$H(O,P) = \sqrt{\frac{1}{2}\sum_i (\sqrt{o_i} - \sqrt{p_i})^2}.$$

For intuitive interpretation, we define prediction success $h = 1 - H$, which ranges from 0 to 1, with the value 1 indicating that all points lie on the 1:1 line. Second, we use the Pearson correlation coefficient between ranks (ρ) because it also avoids assessing the fit with respect to the best-fit line (as R^2 does) and has relatively low sensitivity to outliers. Haegeman and Loreau (2008) have suggested using the root mean squared error, but this has no upper bound and therefore makes interpretation less intuitive. Appendix C provides additional details on fit metrics.

Results

Below, we organize our results with respect to the questions posed in *Methods*.

Are CATs predictable along environmental gradients?—The CATs demonstrated remarkable predictability as quantifiers of constraints on local community composition. Fig. 1 shows the regularity of the traits' variation along the gradient for the fynbos data (mean $r = 0.83$); see Shipley et al. (2006: Fig. 1) for similar vineyard results (mean $r = 0.88$).

How sensitive are abundance predictions to deviations from observed CATs?—T_{sm} consistently had the lowest fit (Figs. 2 and 3, Table 2; Appendix D: Figs. D1–7) when used with the uniform prior, indicating EM's sensitivity to CAT prediction (compare to T_{reg}, which uses observed CATs). The lower accuracy for vineyard abundance predictions under T_{sm}, compared to the fynbos, occurred because these were largely driven by only four traits (Supplement: Table S4) and the second most important trait, aboveground vegetative mass, had the worst predicted CATs ($r = 0.70$). Cross validation on the CATs produced similar, but slightly worse, abundance predictions than T_{sm} (Table 2; Appendix D: Tables D2 and D3). The regular variation in CATs along both successional and elevational gradients meant that CATs splines changed minimally when omitting a single community from the fit under cross validation. Notably, the predictions for each local community in the vineyard (Appendix D: Tables D1–3) were not as poor as previously reported (Marks and Muller-Landau 2007).

How well does EM perform with varying species richness?—The highest h values occurred for T_{loc} for both data sets compared to T_{reg} (Table 2; Appendix D: Table D1). T_{reg} predicted dominant species with similar accuracy to T_{loc}; however, intermediate and rare species were less accurately predicted (Fig. 3, Table 2; Appendix D: Table D1). Despite lower accuracy for rare species, they were seldom classified incorrectly as intermediate or dominant species (Fig. 3); among all 704 possible abundances predicted over both data sets, no more than 10 predictions had an absolute error of 10% or greater across all models (Appendix D: Fig. D18). The decreased accuracy resulted from increased trait vector similarity in the regional (as opposed to local) pool of species. In the fynbos, many dominant species, e.g., common proteoids, ericoids, and restioids, had trait vectors similar to those of rare species in the same guilds; thus additional species in T_{reg} can confound predictions.

Fig. 3 shows predictions for varying numbers of species and traits, as summarized by h value surfaces (Appendix D: Fig. D1, using ρ). Both data sets showed a peak in h at intermediate richness. At richness values above this peak, trait vector similarity confounded predictions. Below this peak, truncating the regional species pool removed species with unique trait vectors that were necessary to fulfill the constraints without substantially altering the RSA. For example, a species with traits very similar to the CATs would be predicted to have high abundance. But if this species were removed from the model, a linear combination of the remaining species' trait vectors would be necessary to fill the void, and these linear combinations could contort RSA predictions, particularly if the omitted trait vector were very dissimilar to any others.

How does trait vector similarity affect prediction accuracy?—We found a positive relationship between mean pairwise dissimilarity among species in a local community and the prediction accuracy under T_{loc} (Appendix D: Fig. D21; for the vineyard, slope = 0.41, $R^2 = 0.07$; for the fynbos, slope = 1.1, $R^2 = 0.51$). This means that communities with more similarity among trait vectors were harder to predict with EM. Trait vector similarity also decreased accuracy of species-level predictions. Absolute prediction error for each species increased when nearest neighbor distance (in trait space) decreased. We found a negative relationship between the change in absolute prediction error and the change in nearest neighbor distance (Appendix D: Fig. D22; for the vineyard, slope = −0.03, $R^2 = 0.10$; for the fynbos, slope = −0.03, $R^2 = 0.25$), confirming our expectation that increasing similarity in trait space decreases prediction accuracy. Scatter about the regression lines is not surprising because some species may be relatively distinct in trait space, so an increase in trait vector

TABLE 2. Summary of model fit under different tests for the vineyard and fynbos data sets.

Test	EM	Permutation test		Perturbation test, median and 99% CI		
		Median and 99% CI	Null > true	Recalculate CATs, 5% error	Observed CATs, 5% error	Observed CATs, 20% error
A) Vineyard						
T_{loc}	0.92	0.88 (0.83–0.94)	7%	0.92 (0.91–0.92)	0.86 (0.82–0.88)	0.66 (0.60–0.73)
T_{reg}	0.68	0.61 (0.52–0.69)	1.6%	0.67 (0.66–0.69)	0.65 (0.60–0.68)	0.50 (0.31–0.60)
T_{sm}	0.46	0.14 (0.09–0.32)	0%	0.45 (0.43–0.46)	0.45 (0.43–0.46)	0.39 (0.23–0.44)
T_{sm}, regional prior	0.46	0.15 (0.08–0.32)	0%	0.45 (0.44–0.46)	0.44 (0.42–0.45)	0.39 (0.23–0.46)
T_{sm}, subreg. prior	0.48	0.17 (0.11–0.38)	0%	0.48 (0.46–0.48)	0.46 (0.44–0.48)	0.40 (0.21–0.47)
B) Fynbos						
T_{loc}	0.76	0.65 (0.59–0.71)	0%	0.76 (0.75–0.77)	0.67 (0.66–0.70)	0.73 (0.66–0.75)
T_{reg}	0.55	0.46 (0.42–0.51)	0%	0.56 (0.54–0.56)	0.56 (0.54–0.56)	0.54 (0.52–0.56)
T_{sm}	0.51	0.46 (0.42–0.55)	0%	0.51 (0.51–0.52)	0.50 (0.50–0.51)	0.51 (0.49–0.52)
T_{sm}, regional prior	0.58	0.53 (0.50–0.57)	0%	0.57 (0.57–0.59)	0.57 (0.57–0.58)	0.58 (0.54–0.59)
T_{sm}, subreg. prior	0.63	0.62 (0.61–0.63)	6.6%	0.62 (0.61–0.63)	0.62 (0.61–0.63)	0.58 (0.61–0.63)

Notes: Fit of predicted and observed relative abundance was measured with $h = 1 -$ Hellinger divergence (median and 99% CI). T_{loc}, T_{reg}, and T_{sm} refer, respectively, to tests with local and regional species pools and regional pools using smoothed splines; T_{sm} used predicted CATs, whereas T_{loc} and T_{reg} use observed CATs. For the permutation tests, the percentage of null models better than the true model is given. For the perturbation tests, we considered two cases: (1) observed CATs were used with permuted trait matrix, and (2) CATs were recalculated based on the permuted trait matrix. Regression model R1 regresses abundance on raw trait values; R2 uses the squared difference between each species' traits and CATs for the each local community. For the perturbation tests, 5% (or 20%) error indicates that perturbed traits were drawn from a uniform distribution ranging from 5% (or 20%) below to 5% (or 20%) above observed values for each trait. Regression model R1 has entries of "NA" (not applicable) for tests incorporating CATs because R1 does not use CATs. Regression models have "NA" entries for nonuniform priors because we use maximum-likelihood models that do not incorporate this information.

similarity may not appreciably change EM's ability to distinguish that species from others.

How much trait information is necessary or sufficient? Which traits provide the most predictive power? Do quantitative vs. categorical traits matter?—Lagrange multipliers for each trait are shown in the Supplement: Table S4. The traits with the most predictive power (in order, averaged across all models) in the vineyard were stem mass, aboveground vegetative mass, and seed number, whereas in the fynbos these were succulence, flowering duration, and stem length/stem basal diameter. Prediction accuracy increased using more traits (Fig. 3), but the vineyard surface was more plateau-like for more than four traits. This occurred because four traits primarily drove model fit (Appendix D: Table D2). Categorical traits were comparable in predictive value to quantitative traits (Appendix D: Table D5).

How do different priors affect prediction accuracy? Can priors be used to model dispersal limitation effects?—Predictions improved in all cases with informative empirical priors that reflect null expectations about regional or subregional abundance. We focus on the predictions for T_{sm} because this test corresponds to practical application of EM, although predictions for all three T_{loc} and T_{reg} were also improved using informative priors (data not shown). In all cases, subregional priors provided better predictions than regional priors, which were better than uniform priors. The vineyard predictions improved only minimally, whereas the fynbos data were much better predicted, particularly for subregional priors (Fig. 3).

Are predictions sensitive to intraspecific trait variation or measurement error at the species level?—EM was robust to perturbing the measured trait values, as h values never decreased by more than 0.10 under perturbation by 5% (Table 2; see Supplement: Table S2 for results using ρ). Perturbation by 20% led to more substantial decreases in the vineyard and demonstrated that inaccurate measurements of traits or high intraspecific variability can lead to substantial degradation of abundance predictions. In the fynbos, h values decreased less, but this was largely due to the fact that we did not perturb the four categorical traits, and two of these (perenniality and succulence) had among the largest Lagrange multipliers. Intraspecific trait variation is clearly a major obstacle for species-level predictions along gradients, but our predictions appear to be robust to at least some level of intraspecific variation or measurement error in traits (<5%).

Does EM perform better than null models?—All EM predictions were better than at least 98.4% of null models (Table 2; Appendix D: Table D1) except for T_{loc} in the vineyard and T_{sm} with the subregional prior for the fynbos. Null models were more accurate in the former case because there were relatively few degrees of freedom in the low-richness vineyards. In the latter case, the subregional priors had a particularly strong influence on predictions because the prior RSAs were sufficiently similar to the true RSAs.

Does EM perform better than regression models, using the same information?—For both regression models, R1 and R2, T_{loc} made the best predictions whereas T_{reg} and T_{sm} were comparable to one another (Table 2; Appendix D: Figs. D19 and D20). EM performed better than, or comparable to, both regression models in all cases (Table 2; Appendix D: Table D1). Note that fitting

TABLE 2. Extended.

Cross validation	Regression model 1	Regression model 2
NA	0.92	0.92
NA	0.58	0.57
0.37	NA	0.57
0.37	NA	NA
0.37	NA	NA
NA	0.77	0.77
NA	0.51	0.51
0.44	NA	0.51
0.49	NA	NA
0.50	NA	NA

regression models required ad hoc assumptions about observed zeroes, unlike EM. Also, because there were more traits than species (only) under T_{loc} for the vineyard data, seven of 12 communities were uniquely determined by the regression models.

DISCUSSION

Functional traits effectively predict community composition

By using a functional approach, EM shows that generality in community ecology emerges from patterns of functional trait distributions along ecological gradients (c.f. McGill et al. 2006). Species' functional traits mediate their interaction with the surrounding community through habitat filtering. Thus, modeling functional patterns (CATs) not only yields a more tractable, lower-dimensional problem than species-level models, but also reflects our mechanistic understanding of species interactions. We show that CATs respond strongly and relatively smoothly to ecological gradients and that their usefulness as predictors is robust to realistic levels of intraspecific variation and measurement error. Additionally, suites of CATs demonstrate a desirable amount of sensitivity in their ability to differentiate communities; abundance predictions are robust to minor variations in CATs, but differences among CATs along gradients successfully distinguish communities. EM links species-level predictions back to the functional trait patterns that are modeled by CATs.

The generality of EM hinges on the predictability of CAT variation along ecological gradients. If traits are important to habitat filtering, and CATs describe the biological response to environmental conditions, they should vary regularly along environmental gradients. Irregular variation in CATs could reflect under-sampling, traits under weak selection, missing environmental variables defining the gradient, or inadequacy of CATs to characterize niche constraints. Shipley et al. (2006) demonstrated good correlation of CATs with spline fits for the vineyard data, and we have shown that CATs for the more diverse fynbos have similarly good fits (Fig. 1). The fact that some variation in CATs about the trend line remains is not surprising, because we do not necessarily expect strong selection on all traits at all points along a gradient.

In spite of the strong evidence that CATs reflect habitat filtering at some level, questions remain. First, CAT values are heavily influenced by dominant species. Dominant species are typically the most important in terms of community-level processes, so CATs reasonably capture broadscale changes in communities. But as Westoby and Wright (2006) have pointed out, a great deal of trait variation exists within communities, and CATs do not capture subtle variation among rare species. Considering trait variance to account for niche differentiation within communities, or using CATs that are not weighted by abundance (to increase the influence of rare species on constraints), could serve to address these concerns. Second, the abundances used to calculate CATs contain all processes that affect abundance, not just niche-based processes. Niche-based processes should dominate a community if the CATs are used to infer something about habitat suitability. We postulate that when the CATs follow a trend, as in our data (Fig. 1), they indicate habitat filtering as a dominant mechanism driving abundance patterns.

The predictions of our models are sensitive to the CATs; the difference between T_{reg}, where observed CATs are used, and T_{sm}, where predicted CATs are used, indicates that minor deviations in the CATs lead to different predicted abundances (Figs. 2 and 3). We anticipate that more robust abundance predictions will result from splines fit with more communities; regressions using 8–12 points can be relatively sensitive to sample bias. Nonetheless, from the two contrasting systems that we studied, we conclude that CAT variation along the gradient captures the influence of environmental filtering on local community composition and the constraints they impose are sufficient to differentiate local communities.

Our results using spline-predicted CATs (T_{sm}) and cross validation address the robustness of predicted RSAs. Unlike Shipley et al. (2006), we found that predicted abundance using predicted CATs with both data sets leads to decreased prediction accuracy (T_{sm}, Fig. 2; compare to $r = 0.97$ and $r = 0.96$ reported by Shipley et al. 2006). However, Shipley et al. smoothed both the empirical and predicted abundances, whereas we did not. Their success with predicted CATs is largely due to the common shift of high-influence empirical and predicted abundances under smoothing. However, the reduced accuracy of T_{sm} compared to T_{reg} is not detrimental; this simply indicates that the gradient must be well defined. To demonstrate this, note that our fynbos predictions showed much higher similarity between the accuracy of T_{reg} and T_{sm} (Fig. 2) because the traits with the largest Lagrange multipliers (peren-

niality, succulence, height) were among the best-predicted along the gradient.

Trait vector similarity and prediction accuracy

One challenge in EM is to understand the factors affecting prediction accuracy. A key to predictive accuracy lies in the uniqueness of the species' trait vectors. To understand how trait vectors affect predictions, consider a number of species with similar traits (e.g., belonging to a functional group) that coexist in a regional community. EM predicts a similar abundance for each species. For such a group, a large feasible set exists, which includes all possible combinations of abundances. In contrast, if a species pool with distinct traits were subjected to the same CATs, they would experience stronger constraints. This contrast can be observed by comparing T_{loc}, which contains the minimum number of species to exactly satisfy the constraints, to T_{reg}, where additional species with potentially similar trait vectors are included (Fig. 2). If species' true functional similarity is well characterized by the measured traits, dissimilarity among traits reduces the volume of the feasible set. However, interpretations of trait vector similarity must be made cautiously because many species may seem to be similar based on the measured traits, while their unmeasured traits may differ greatly.

In the context of what is known about fynbos ecology, we hypothesize that the trait vector similarity in the fynbos data represents actual functional redundancy. Co-occurrence of ecologically similar species in Cape Floristic shrubland communities, represented by proteoids (large woody shrubs similar to the Proteaceae), ericoids (small woody shrubs similar to the Ericaceae), and restioids (graminoids similar to the Restionaceae), has been widely noted at different spatial scales (e.g., Cowling et al. 1994). The contrast in prediction success between the vineyard and fynbos data can be explained in terms of the greater trait vector similarity of the fynbos (Appendix D: Fig. D21). Although we expect the fynbos to have higher functional redundancy, the much larger sample size in the local communities of the fynbos probably contributes to an increase in the diversity sampled compared to the vineyard data (respectively, 3–7 samples of 5×10 m each per local community, vs. one 0.5×0.5 m sample). This higher local diversity, coupled with trait vector similarity, produces the poorer predictions of T_{loc} in the fynbos.

The apparent widespread functional redundancy in some systems, like the fynbos, may represent an insurmountable barrier to accurate species level predictions. Although our predictions probably would improve with additional traits and more complete sampling along the elevational gradient, we believe that true functional redundancy sets an upper bound on the accuracy of niche-based predictions. It may be the case that constraints operate only on functional groups and that abundance patterns within functional groups are dominated by stochastic processes. Thus the best-case scenario for a deterministic model is not perfect predictions, but one that provides some reasonable way of averaging over all the possibilities introduced by stochasticity in local community samples. EM solves this problem by providing a robust prediction method that identifies the only unbiased RSA with respect to the constraints. Indeed, the difficulty in predicting the abundance of particular species is anticipated by a functional approach to community ecology because it asserts that traits, and not species, show the strongest correlation with environmental conditions. Hence the CATs capture broad patterns in traits and EM links these patterns back to the individual species level to the extent possible.

Using informative priors with EM

One possibility for improving predictions based on prior expectations is to use nonuniform priors that are either informative or based on null model expectations (see also Shipley 2009*c*, Sonnier et al. 2010). Using the regional RSA as the prior represents the null hypothesis that the local RSA is expected to be equivalent to the regional RSA in the absence of constraints. Priors can also be interpreted as informative if dispersal limitation is a factor; one can construct the prior RSA from a subset of sites proximate to the site of interest, as we have done with subregional priors. Informative priors can also reflect unmeasured environmental variation (not encompassed in the gradient) that might preferentially influence the abundance of some species.

Our interpretation of using regional and subregional priors is supported by the contrasting results from the two data sets we analyzed and their respective ecology. The vineyard data were better predicted by the regional prior than the uniform prior, providing further evidence for this system that regional abundance substantially affects local community abundance (Lavorel and Lebreton 1992). The plots for the vineyard data are within a few kilometers of each other, dispersal limitation is not known to be particularly important in this successional system, and local composition is more closely related to regional composition than to local seed banks (Lavorel and Lebreton 1992), so its not surprising that subregional priors did not improve on predictions from regional priors for this data set. In contrast, the fynbos data are collected over a range of >50 km and dispersal limitation is a well-known factor in fynbos (Cowling et al. 1997, Latimer et al. 2009); hence, we expect the subregional prior to outperform the regional prior in this system. Thus, for ecological inference, we advocate using EM with T_{sm} and exploring suitable priors that capture regional abundance patterns at the appropriate spatial scale.

Guidelines for using EM for ecological data

Important decisions and assumptions must be made that may be unique to EM when preparing a data set for

analysis. In the Supplement, we provide example R code (R Development Core Team 2009) that can be applied to an arbitrary data set. First, the species pool is truncated to focus on only the most regionally abundant species, following other studies that focus on community-level function (Garnier et al. 2004, Baraloto et al. 2010). Due to this truncation, we point out that EM is most suitable for analyzing broad functional trait and abundance patterns in relatively dominant species. Rare species present a sampling problem (Preston's veil line; Chisholm 2007); because they contribute very little to CATs (Lavorel et al. 2008), we would not expect EM to tell us much more than broadly distinguishing between rare and dominant species. Further, truncation increases the signal-to-noise ratio in community-level patterns of response to habitat filtering by removing rare species whose abundance may depend primarily on stochastic factors (e.g., lottery dynamics). What is the "right" number of species? Roxburgh and Mokany (2007) have explored this problem using simulations to determine a sufficient number of degrees of freedom to make nontrivial predictions; both data sets that we examined have sufficient degrees of freedom to avoid spurious predictions (confirmed by null models; Table 2). We chose species above a threshold of 0.05% regional abundance in the fynbos to represent abundant species, based on a distinct drop-off in abundance for species below this level.

Which are the "right" traits to use? This question pervades the functional trait literature (e.g., Garnier et al. 2004, Petchey and Gaston 2006). To provide information on habitat filtering, traits must vary along the gradient of interest. Starting with a large pool of traits, one can omit traits with the lowest Lagrange multiplier values. Given traits, one can test for the significance of the predictions using the permutation tests just discussed. Finally, what is the consequence of omitting important traits? Information theory postulates that other RSAs will be more likely than the EM-predicted RSA if important trait information is missing (Shipley et al. 2007). We suggest that the importance of missing traits can be ascertained from the accuracy of T_{loc}; if the model fits poorly when the observed CATs are supplied and the effects of trait vector similarity are minimized (compared to T_{reg}), there should be additional traits that lead to variation in abundance.

Finally, we draw attention to the two sources of criticism in EM: circularity (Marks and Muller-Landau 2007) and the number of degrees of freedom (Haegeman and Loreau 2008). It is important to note that the tests we propose focus on model-fitting diagnostics that use observed CATs directly (T_{loc}, T_{reg}) or indirectly (T_{sm}); i.e., they evaluate model fit rather than out-of-sample prediction. However, noncircular applications are possible when CATs are predicted that are *not* used to fit CAT splines along gradients. We demonstrate this using cross validation on CATs and point out that use of holdout data avoids the circularity issue (Table 2). We advocate maximizing the number of degrees of freedom by minimizing the number of traits necessary to adequately explain abundance patterns in a particular application. One can test for sufficient degrees of freedom using the permutation tests previously described; excessively constrained models will perform no better than null models. In general, our models have an acceptable number of degrees of freedom compared to traditional models (e.g., regression), but many fewer than found in many physics applications of MEF (Haegeman and Loreau 2008). By subtracting the number of constraints (number of traits plus normalization) from the number of species, we find that the vineyard data had 21 degrees of freedom and the fynbos had 32, and they are unequivocally not over-determined under T_{reg} and T_{sm} (cf. Haegeman and Loreau 2008).

Conclusions and future directions

While we have demonstrated that EM can provide ecological insight and reasonably accurate predictions, some questions remain to be answered before EM can become an accepted ecological tool. First, a better understanding of CATs and their ability to quantify community-level selection is necessary. How stable are CATs among communities under the same environmental conditions (cf. Acosta et al. 2008)? Is an average trait value adequate to define habitat filtering across the spectrum of local communities? Or is intracommunity trait variability comparable to intercommunity variability (i.e., large trait variance among co-occurring species), as might be expected under niche differentiation (cf. Westoby and Wright 2006)? The answers to these questions will depend on the spatial scale of the samples, but our results suggest that, despite noise and process variability, CATs can provide useful predictions of community composition across typical ecological gradients.

The fact that we have obtained meaningful results using a handful of easily observed and scored traits that do not necessarily encompass all attributes of a plant's biology is testimony to the potential utility of this approach to predict community abundance patterns along ecological gradients. The challenge remains, however, to incorporate EM into an explicitly spatial model with historical and/or stochastic processes that account for recruitment limitation. Informative priors appear to be a step in the right direction. He (2010) has recently noted a connection between EM and multinomial logistic regression; formulating a regression model in a Bayesian framework would readily incorporate informative priors and explicit spatial effects to provide a competing model for EM and represents an important avenue for future research.

EM provides a first step toward the goal of finding general rules of community ecology using a functional approach along ecological gradients (c.f. McGill et al. 2006). EM's novel ability to synthesize functional trait distributions along environmental gradients and to

generate testable predictions provides a much needed quantitative framework to bridge the gap between functional traits and community patterns.

Acknowledgments

Helpful comments made by Bart Haegeman, Kent Holsinger, Bill Shipley, Adam Wilson, and anonymous reviewers were valuable in improving the manuscript. C. Merow was funded by two grants from the Center for Environmental Sciences and Engineering (CESE) at the University of Connecticut and NSF grants DEB 0516320 and OISE 0623341 to J. A. Silander.

Literature Cited

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

Acosta, B., L. Sanchez-Jardon, A. del Pozo, E. Garcia-Ibanez, M. A. Casado, J. Montalvo, and F. D. Pineda. 2008. Grassland species composition and morpho-functional traits along an altitudinal gradient in a Mediterranean environment: relationship with soil water availability and evaporative dynamic. Acta Oecologica International Journal of Ecology 34:26–37.

Albert, C. H., W. Thuiller, N. G. Yoccoz, A. Soudant, F. Boucher, P. Saccone, and S. Lavorel. 2010. Intraspecific functional variability: extent, structure and sources of variation. Journal of Ecology 98:604–613.

Ali, M. S., and D. Silvey. 1966. A general class of coefficients of divergence of one distribution from another. Journal of the Royal Statistical Society B 28:131–140.

Baraloto, C., C. E. T. Paine, S. Patino, D. Bonal, B. Hérault, and J. Chave. 2010. Functional trait variation and sampling strategies in species-rich plant communities. Functional Ecology 24:208–216.

Chisholm, R. A. 2007. Sampling species abundance distributions: resolving the veil-line debate. Journal of Theoretical Biology 247:600–607.

Cowling, R. M., P. J. Mustart, H. Laurie, and M. B. Richards. 1994. Species diversity, functional diversity and functional redundancy in fynbos communities. South African Journal of Science 90:333–337.

Cowling, R. M., D. M. Richardson, and P. J. Mustart. 1997. Fynbos. Pages 99–130 *in* R. M. Cowling, D. M. Richardson, and S. M. Pierce, editors. Vegetation of Southern Africa. Cambridge University Press, Cambridge, UK.

Dewar, R. C., and A. Porteì. 2008. Statistical mechanics unifies different ecological patterns. Journal of Theoretical Biology 251:389–403.

Diaz, S., M. Cabido, and F. Casanoves. 1998. Plant functional traits and environmental filters at a regional scale. Journal of Vegetation Science 9:113–122.

Garnier, E., J. Cortez, G. Billés, M.-L. Navas, C. Roumet, M. Debussche, G. Laurent, A. Blanchard, D. Aubry, A. Bellmann, C. Neill, and J.-P. Toussaint. 2004. Plant functional markers capture ecosystem properties during secondary succession. Ecology 85:2630–2737.

Goldblatt, P., and J. C. Manning. 2000. Cape plants: a conspectus of the Cape flora of South Africa. National Botanical Institute and Missouri Botanical Garden, Pretoria, South Africa and St. Louis, Missouri, USA.

Haegeman, B., and M. Loreau. 2008. Limitations of entropy maximization in ecology. Oikos 117:1700–1710.

Haegeman, B., and M. Loreau. 2009. Trivial and non-trivial applications of entropy maximization in ecology: A reply to Shipley. Oikos 118:1270–1278.

Harte, J., A. B. Smith, and D. Storch. 2009. Biodiversity scales from plots to biomes with a universal species–area curve. Ecology Letters 12:789–797.

Harte, J., T. Zillio, E. Conlisk, and A. B. Smith. 2008. Maximum entropy and the state-variable approach to macroecology. Ecology 89:2700–2711.

He, F. 2010. Maximum entropy, logistic regression, and species abundance. Oikos 119:578–582.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.

Jaynes, E. T. 1957. Information theory and statistical mechanics. Physical Review 106:620–630.

Jaynes, E. T. 2003. Probability theory: the logic of science. Cambridge University Press, Cambridge, UK.

Keddy, P. A. 1992. Assembly and response rules: two goals for predictive community ecology. Journal of Vegetation Science 3:157–164.

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.

Kvalseth, T. O. 1985. Cautionary note about R^2. American Statistician 39:279–285.

Latimer, A. M., J. A. Silander, A. G. Rebelo, and G. F. Midgley. 2009. Experimental biogeography: the role of environmental gradients in high geographic diversity of Cape Proteaceae. Oecologia 160:151–162.

Lavorel, S., K. Grigulis, S. McIntyre, D. Garden, N. Williams, J. Dorrough, S. Berman, F. Quétier, A. Thébault, and A. Bonis. 2008. Assessing functional diversity in the field—methodology matters! Functional Ecology 22:134–147.

Lavorel, S., and J. D. Lebreton. 1992. Evidence for lottery recruitment in Mediterranean old fields. Journal of Vegetation Science 3:91–100.

Manly, B. F. J. 2006. Randomization and Monte Carlo methods in biology. Third edition. Chapman and Hall, London, UK.

Marks, C. O., and H. C. Muller-Landau. 2007. Comment on From plant traits to plant communities: a statistical mechanistic approach to biodiversity. Science 316:1425.

McGill, B. 2003. Strong and weak tests of macroecological theory. Oikos 102:679–685.

McGill, B. J., B. J. Enquist, E. Weiher, and M. Westoby. 2006. Rebuilding community ecology from functional traits. Trends in Ecology and Evolution 21:178–185.

McGill, B. J., et al. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. Ecology Letters 10:995–1015.

McGill, B., and J. C. Nekola. 2010. Mechanisms in macroecology: AWOL or purloined letter? Towards a pragmatic view of mechanism. Oikos 119:591–603.

Petchey, O. L., and K. J. Gaston. 2006. Functional diversity: back to basics and looking forward. Ecology Letters 9:741–758.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231–259.

Preston, F. W. 1948. The commonness and rarity of species. Ecology 29:254–283.

Proches, S., R. M. Cowling, and L. Mucina. 2003. Releve-based species–area curves for the Cape Floristic Region. South African Journal of Science 99:474–476.

Pueyo, S., F. He, and T. Zillio. 2007. The maximum entropy formalism and the idiosyncratic theory of biodiversity. Ecology Letters 11:1017–1028.

R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ⟨http://www.R-project.org⟩

Rebelo, A. G., C. Boucher, N. Helme, L. Mucina, and M. G. Rutherford. 2006. Fynbos biome. Pages 52–219 *in* L. Mucina

and M. C. Rutherford, editors. The vegetation of South Africa, Lesotho and Swaziland. Strelitzia 19. South African National Biodiversity Institute, Pretoria, South Africa.

Roxburgh, S. H., and K. Mokany. 2007. Comment on "From plant traits to plant communities: a statistical mechanistic approach to biodiversity." Science 316:1425.

Schumaker, L. L. 2007. Spline functions: basic theory. Cambridge University Press, New York, New York, USA.

Shipley, B. 2009*a*. Limitations of entropy maximization in ecology: a reply to Haegeman and Loreau. Oikos 118:152–159.

Shipley, B. 2009*b*. Trivial and non-trivial applications of entropy maximization in ecology: Shipley's reply. Oikos 118:1279–1280.

Shipley, B. 2009*c*. From plant traits to vegetation structure: chance and selection in the assembly of ecological communities. Cambridge University Press, Cambridge, UK.

Shipley, B., D. Vile, and É. Garnier. 2006. From plant traits to plant communities: a statistical mechanistic approach to biodiversity. Science 314:812–814.

Shipley, B., D. Vile, and É. Garnier. 2007. Reply to comments on "From plant traits to plant communities: a statistical mechanistic approach to biodiversity." Science 316:1425.

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 Vegetation Science 21:318–331.

Volkov, I., J. R. Banavar, S. P. Hubbell, and A. Maritan. 2009. Inferring species interactions in tropical forests. Proceedings of the National Academy of Sciences USA 106:13854–13859.

Westoby, M., and I. J. Wright. 2006. Land-plant ecology on the basis of functional traits. Trends in Ecology and Evolution 21:261–268.

Wiegart, R. G. 1962. The selection of an optimum quadrat size for sampling the standing crop of grasses and forbs. Ecology 43:125–129.

APPENDIX A

Model formulation (*Ecological Archives* E092-127-A1).

APPENDIX B

A description of the data collection (*Ecological Archives* E092-127-A2).

APPENDIX C

Assessing model fit (*Ecological Archives* E092-127-A3).

APPENDIX D

Additional analyses including alternate fit metrics, local community predictions, informative prior predictions, regression models, trait vector similarity, and cross validation (*Ecological Archives* E092-127-A4).

SUPPLEMENT

R code and fynbos data set (*Ecological Archives* E092-127-S1).