Inferential Tests and Modelling of Functional Trait Convergence along Environmental Gradients

Chaitanya Joshi[1,3], Paul T. Brown[1], and Daniel C. Laughlin[2]

[1]Department of Statistics, University of Waikato, Hamilton, New Zealand
[2]Department of Biological Sciences, University of Waikato, Hamilton, NZ
[3]Corresponding author: Chaitanya Joshi, email cjoshi@waikato.ac.nz

Abstract

The motivation for this paper comes from a recent study which indicated that the influence of environmental filtering should increase with decreasing soil fertility, based on the premise that individuals will employ a resource-retentive strategy in a less productive ecosystem. Mean annual temperature (MAT) is one indicator of the productivity of the ecosystem. We aim to build a more accurate model of environmental filter and want to statistically test whether the environmental filter is stronger when the MAT is lower compared to when it is higher. Our findings throw an interesting insight into how the trait variability changes as a function of MAT and how it could be better modelled.

Keywords: Bayes factors, Likelihood ratio test, Environmental filter, Markov Chain Monte Carlo, Trait based community assembly.

1 Introduction

Community assembly theory suggests that individuals are sorted along environmental gradients as their functional traits influence their fitness and performance (Keddy (1992)). Species with similar functional traits will more likely be found in similar environments, leading to convergence of trait values within communities (Shipley (2010)). The motivation for this paper comes from a recent study by Mason et al. (2012), which indicated that the influence of environmental filtering should increase with decreasing soil fertility, based on the premise that individuals will employ a resource-retentive strategy in a less productive ecosystem, and that small differences in competitive ability will not have disproportionate effects on the outcome of competition (Lambers et al. (2008)). Mean annual temperature (MAT) is one indicator of the productivity of the ecosystem. We aim to build a more accurate model of environmental filter and want to statistically test whether the environmental filter is stronger when the MAT is lower compared to when it is higher.

2 Data and Models

We re-analyse the data used by Laughlin et al. (2012). They measured three different functional traits - bark thickness, specific leaf area (SLA) and wood density. The three functional traits were measured on individual trees from nine different species across several sites in Arizona, USA. The sites had a 10°C range in MAT. The traits were chosen to represent key spectrums of plant strategies, and are not strongly correlated with each other. Laughlin et al. (2012) modelled the environmental filters (for each trait) by using generalised linear models with a log-link function and polynomial regression
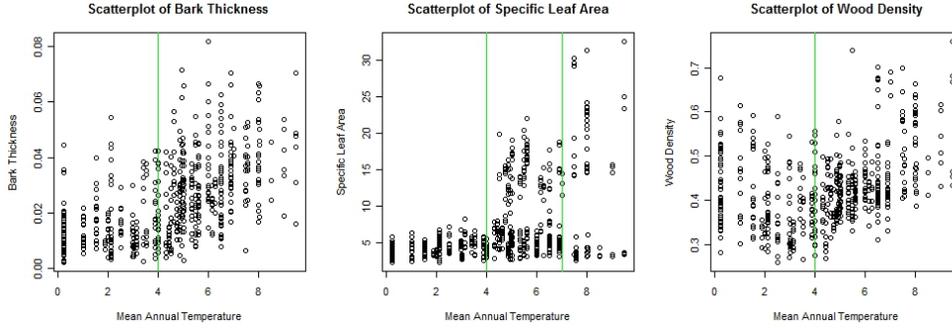
Figure 1: Scatterplots for each trait versus Mean Annual Temperature, vertical lines show the partitions for each trait.

equations with *constant* variance. The models of the three functional traits were as follows:

$$
\begin{aligned}
\log(bark) &= \beta_0 + \beta_1 t + \beta_2 t^2 + \epsilon \\
\log(SLA) &= \beta_0 + \beta_1 t + \beta_2 t^2 + \epsilon \\
\log(wood) &= \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \epsilon
\end{aligned}
$$

where, t denotes the MAT, and ϵ was assumed to follow a normal distribution, with $E(\epsilon) = 0$ and $Var(\epsilon) = \sigma^2 \mathbf{I}$. The log-link function was used to ensure the prediced trait values were positive, and the polynomial equations were used to ensure the curved relationship of each trait versus temperature were accounted for.

A stronger environmental filter implies that the variance will be smaller. If the theory is true then the variability in the trait values should be significantly smaller for lower values of MAT than for the higher values. We test the theory by partitioning the data and refitting the above models assuming *unequal* variances for each partition. The bark and wood data were partitioned at 4°C to distinguish between lower and higher MAT. SLA was partitioned at 4°C and 7°C . We take a conservative approach and test weather the variances are equal or not. Thus, we have the following two-sided hypotheses for the three traits:

$$
\mathcal{H}_0 : \sigma_1{}^2 = \sigma_2{}^2 \quad \text{versus} \quad \mathcal{H}_1 : \sigma_1{}^2 \neq \sigma_2{}^2 \quad \text{for bark and wood,}
$$
$$
\text{and} \quad \mathcal{H}_0 : \sigma_1{}^2 = \sigma_2{}^2 = \sigma_3{}^2 \quad \text{versus} \quad \mathcal{H}_1 : \sigma_1{}^2 \neq \sigma_2{}^2 \neq \sigma_3{}^2 \quad \text{for SLA,}
$$

where σ_1^2 denotes the variance for the lowest partition (MAT < 4°C). Testing these hypotheses involves model fitting under each of the hypotheses.

3   Statistical Inference

Parameter estimation under $\mathcal{H}_0$ is straightforward under both the *classical* as well as the Bayesian paradigm. Under $\mathcal{H}_1$, we first take the Bayesian approach and employ an MCMC based method, namely, a blockwise MH algorithm with random walk proposals. Flat priors were used for objective estimation. MCMC chains converged satisfactorily. Under $\mathcal{H}_1$, no closed form analytical solution exists for the MLE, however, iterative procedures have been proposed. We find the MLE using the two-stage estimator of Rao and Fuller (1978). Both these methods yield nearly identical estimates (Table 1).

Table 1: Parameter estimates using the two-stage estimator (2st) and the posterior means (MCMC) along with the 95% credible intervals for the posteriors. Note these are in the natural log-scale.

Data	Method	b0_hat	b1_hat	b2_hat	b3_hat	sg1_hat	sg2_hat	sg3_hat
Bark	2st	-4.621966	0.140814	0.003518	NA	0.654295	0.517481	NA
	MCMC	-4.620872	0.140342	0.003577	NA	0.660169	0.520704	NA
	credUp	-4.455600	0.214520	0.011280	NA	0.732750	0.561090	NA
	credLo	-4.790211	0.067510	-0.004480	NA	0.598660	0.484750	NA
SLA	2st	1.321840	0.037800	0.009762	NA	0.253839	0.581673	0.897397
	MCMC	1.323722	0.035733	0.009451	NA	0.255802	0.587016	0.917883
	credUp	1.399270	0.089070	0.016660	NA	0.284260	0.642420	1.10292
	credLo	1.246950	-0.016050	0.002120	NA	0.230600	0.536950	0.77407
Wood	2st	-0.784908	-0.153418	0.034251	-0.001664	0.190791	0.149960	NA
	MCMC	-0.784568	-0.153109	0.034124	-0.001655	0.192113	0.150766	NA
	credUp	-0.767110	-0.142700	0.036220	-0.001427	0.212751	0.162270	NA
	credLo	-0.802410	-0.162210	0.031390	-0.001857	0.173866	0.140140	NA

The results in Table 1 show that for all traits, the standard deviations had non-overlapping 95% credible intervals indicating that $P(\mathcal{H}_0|data) < 0.05$. This would suggest that the variances are significantly different in each partition of MAT for each trait. More formally, we test using both the Bayes factors and the likelihood ratio test.

Bayes factors offer a way of evaluating evidence in favour of a null hypothesis (Kass and Raftery (1995)), a Bayesian alternative to the frequentists p-value. However, derivation of the Bayes factor is not obvious when the $\mathcal{H}_0$ or the $\mathcal{H}_1$ correspond to probability 0 events, as in this case, where the $\mathcal{H}_0$ is a singleton. We derive in the appendix, the expression for the Bayes factor:

$$\mathbf{B^{01}} = \frac{\hat{\sigma_c}^{-n} \exp(-\frac{1}{2\sigma_c^2}(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta))}{\frac{1}{(b-a)^\rho} \prod_{i=1}^{\rho} \exp(-\frac{1}{b})(\frac{n_i}{2}!) \sum_{k=0}^{\frac{n_i}{2}} \frac{b^{-k}}{k!} - numerator}, \quad (1)$$

where $\hat{\sigma_c}$ is the MCMC estimate for standard deviation assuming constant variance, ρ is the number of partitions, and a and b are the parameters of the uniform priors used. The resulting Bayes factors (Table 2) for all three models turned out to be extremely small (≈ 0), implying that the data offers almost no evidence in favour of $\mathcal{H}_0$ which agrees with our previous test.

The likelihood ratio is defined as:

$$\lambda = \frac{sup_{\mathcal{H}_0} L(y, \sigma_c^2)}{sup_{\theta \in \Omega} L(y, \Sigma)} \quad (2)$$

The numerator is the likelihood function evaluated under $\mathcal{H}_0$ with constant variance σ_c^2, and denominator evaluates the likelihood function under the parameter space Ω with non-constant variance Σ. This ratio of variances follows a chi-square distribution with $(\rho - 1)$ degrees of freedom. The p-values for each trait (Table 2) confirm that there is extremely strong evidence against $\mathcal{H}_0$, verifying our earlier results.

Table 2: Hypothesis test results for the Bayes factor and likelihood ratio tests.

Data	Bay_fac	reject_H0	LRT	p_val	reject.H0
Bark	0	Yes	13.9648	0.00019	Yes
SLA	0	Yes	185.4304	0.00000	Yes
Wood	0	Yes	15.2808	0.00009	Yes

These results are not surprising since Figure 1 clearly shows that the variances are not constant. Point estimates for SLA confirm that $\sigma_1 < \sigma_2 < \sigma_3$, in line with

theory we set out to test. However, for bark thickness and wood density, the estimates contradict the theory, i.e. they indicate that $\sigma_1 > \sigma_2$. In order to reconfirm that this was indeed the case, we derived Bayes factors (derivation not included) to test the hypotheses:

$$\mathcal{H}_0 : \sigma_1{}^2 = \sigma_2{}^2 \quad \text{versus} \quad \mathcal{H}_1 : \sigma_1{}^2 > \sigma_2{}^2 \quad \text{for bark and wood.}$$

Bayes factors for both the traits turned out to be extremely small (≈ 0) indicating that the data has strong support in favour of the $\mathcal{H}_1 : \sigma_1{}^2 > \sigma_2{}^2$. So what do these findings mean? Do they indicate that the environmental filter is in fact weaker for lower values of MAT?

4 Discussion

A closer look at Figure 1 reveals that firstly, for wood density, the scatter plot, in fact, does show more variation in the lower partition than in the upper and therefore the result ($\sigma_1{}^2 > \sigma_2{}^2$) is not surprising given the data, although it does appear to contradict the premise we set out to test. Secondly, it can be seen that for bark thickness, the data is much more skewed (skewness = 1.4) in the lower partition compared to the upper partition (skewness = 0.6). It is important to note that the data follows a log-normal distribution, and hence the variance will increase with skewness.
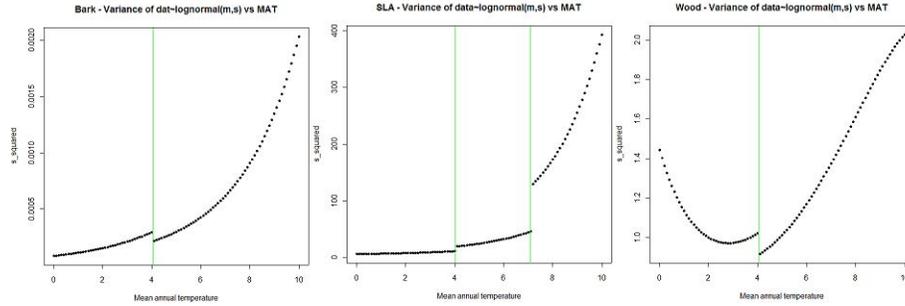

Figure 2: Estimated variance as a function of MAT for each functional trait

Also, note that σ_i's represent the variance of the log transformed data, but it is the variance of the original data (lets denote it by s^2) that should be examined in order to test the theory. If $\log(X) \sim N(\mu, \sigma^2)$ then $var(X) = s^2$ is given by

$$s^2 = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}. \tag{3}$$

Based on the estimates of the regression equation and the σ_i's, it is possible to estimate the variance of the trait data. These estimates of trait variances are plotted against the MAT in Figure 2. It can be seen that the variance of bark thickness in fact increases with increase in MAT and so does the variance of SLA. Variance of wood density, however, shows an increase towards the extreme MAT values on both sides.

The rise in variance of SLA with increasing MAT makes intuitive sense given that only species with low SLA would persist at low temperatures, whereas a greater range of SLA can persist at high temperatures. The rise in the variance of bark thickness with increasing MAT makes intuitive sense given that at low temperatures, fires are much less frequent, and therefore investing resources in thick bark to resist fire is not advantageous. At higher temperatures, a broader range of fire regimes occurs, and

therefore a broader range of bark thickness allows for greater species coexistence across the landscape. The U-shaped response of wood density variance is more difficult to interpret. The high variance at low temperatures is likely due to one species (Pinus aristata) at the highest elevations that has dense wood, since most other species have very light wood at that elevation. The low variance at intermediate temperatures reflects strong filtering for less dense wood at moderately cold temperatures where drought-induced cavitation is less likely than at treeline.

5   Conclusions and Further Work

Prediction of species abundance through the development of mathematical models of trait-based community assembly has seen a renewed interest in recent times (McGill (2006), Shipley et al. (2006), Laughlin et al. (2012)). Accruate modelling of the environmental filter is likely to improve the predictions of such models. This study has given an insight into how the variability of functional trait values changes as a function of MAT. Results for SLA and bark thickness conform to the premise that individuals will employ a resource retentive strategy in a less productive eco-systems. We suspest that the results for wood density are skewed by the presence of one species (Pinus aristata) with unusual characteristics.

It is important to characterise the environmental filter for a wide range of trait values and against several environmental gradients. Figure 2 indicates that modelling the variance as a continuous process (instead of partitioning) is likely to result in more realistic modelling of the environmental filter. This possibility needs to be further explored.

Appendix: Derivation of the Bayes factor

The null hypothesis is a singleton, i.e. of measure zero under any continuous distribution. The prior distribution is therefore defined as: $\pi(\sigma_1, \sigma_2) = \rho_0 \times I_{\Omega_0}(\sigma_1, \sigma_2) + (1 - \rho_0) \times g(\sigma_1, \sigma_2)$, where $\rho_0 = 0.5$ and $g(.,.)$ is the uniform distribution over $(a, b] \times (a, b]$. This ensures that $\int \pi(\sigma_1, \sigma_2) \, d\sigma_1 \, d\sigma_2 = 1$. After simplification, the Bayes factor B^{01} reduces to:

$$\mathbf{B^{01}} = \frac{f(x, \sigma_c)}{\int_{\Omega_1} f(x, \sigma_1, \sigma_2) g(\sigma_1, \sigma_2) \, d\sigma_1 \, d\sigma_2}, \tag{4}$$

where, $f(x, \sigma_c) = (\sqrt{2\pi}\sigma_c)^{-n} \exp(\frac{-\sum_{i=1}^{n}(x_i - \mu)^2}{2(\sigma_c)^2})$, and the denominator is

$$\frac{(\sqrt{2\pi})^{-n}}{(b-a)^\rho} \int_{1/b}^{1/a} \nu_1^{n_1/2} \exp(\frac{-\nu_1 \sum_{i=1}^{n_1}(x_i - \mu)^2}{2}) \, d\nu_1 \times$$

$$\int_{1/b}^{1/a} \nu_2^{n_2/2} \exp(\frac{-\nu_2 \sum_{i=1}^{n_2}(x_i - \mu)^2}{2}) \, d\nu_2,$$

where the integral is expressed in terms of precision $\nu_i = \sigma_i^{-2}$. The intergrands in the denominator are incomplete gamma functions and for $n_1/2$ and $n_2/2$, both integers, the integrals have a closed form solution. Using these along with some simplification yeilds Equation(1).

References

Kass, R. and A. Raftery (1995). Bayes factors. *Journal of the American Statistical Association 90*, 773–795.

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

Lambers, H., J. Raven, G. Shaver, and S. Smith (2008). Plant nutrition acquisition strategies change with soil age. *Trends in Ecology and Evolution 23*, 95–103.

Laughlin, D., C. Joshi, P. van Bodegom, Z. Bastow, and P. Fule (2012). A predictive model of community assembly that incorporates intraspecific trait variation. *Ecology Letters 15*, 1291–1299.

Mason, N., S. Richardson, D. Peltzer, F. de Bello, D. Wardle, and R. Allen (2012). Changes in coexistence mechanisms along a long-term chronosequence revealed by functional trait diversity. *Journal of Ecology 100*, 678–689.

McGill, B. (2006). A renaissance in the study of abundance. *Science 314*, 770–772.

Rao, J. and W. Fuller (1978). Estimation for a linear regression model with unknown diagonal covariance matrix. *The Annals of Statistics 6*, 1149–1158.

Shipley, B. (2010). *From plant traits to vegetation structure: chance and selection in the assembly of ecological communities.* Cambridge, UK: Cambridge University Press.

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