Supplement 3: Model fitting and results

Peter A. Vesk, William K. Morris, Will C. Neal,

Karel Mokany & Laura J. Pollock

January 2020

Reference model building and fitting

We used the same approach to trait-environment modelling as in (Pollock *et al.* 2012) and (Pollock *et al.* 2018). We employ generalised linear mixed models (GLMM) with intercepts and slopes varying by taxon, and fixed effects for traits modulating those slopes. These are variously known as hierarchical models or multilevel models (Miller *et al.* 2019). We fit models of the form:

$$
\begin{aligned}
\Pr(Y_{ij} = 1) &= \text{logit}^{-1}(X_i \beta_j) \\
\beta_k &\sim \text{N}(\mu_k, \Sigma) \\
\mu_{jk} &= \begin{cases} \alpha & : k = 0 \\ Z_j \gamma_k & : k > 0 \end{cases} \\
\alpha, \gamma_{mk} &\sim \text{N}(0, s) \\
\Sigma &\sim \text{Inv-Wishart}\left(v, I\sqrt{\tfrac{v}{2}}\right)
\end{aligned}
\tag{1}
$$

Here the probability that the i^{th} of N sites is occupied by the j^{th} of J taxa ($Y_{ij} = 1$) is the inverse-logit inner product of K environmental gradient values, X_i, and gradient coefficients, β_j. Thus, Y, X and β are N by J, N by K and K by J matrices, respectively. For the first column ($k = 0$) of X, all values are 1, accounting for the gradient intercepts. The rows of the gradient coefficient matrix, β_k, are multivariate-normal distributed with mean vectors, μ_k, and covariance, Σ, such that μ, representing the expected response of taxa to gradients given their traits, has the reverse of the dimensions of β, and Σ is of size K. When $k = 0$, μ_{jk} is equal to α, the overall expected prevalence, so that traits do not influence taxon prevalence. For $k > 0$, μ_{jk}, the expected response of the j^{th} taxon to the k^{th} gradient is the inner product of the j^{th} taxon's M traits, Z_j and M trait-k^{th}-gradient interaction coefficients, γ_k. Thus, Z is a J by M matrix of taxon traits and γ is an M by $K - 1$ matrix of trait-gradient coefficients. The parameter α and the elements of γ are normally distributed around 0 with variance, s. The covariance, Σ, is inverse-Wishart distributed with degrees of freedom, v, and a scale matrix $I\sqrt{\frac{v}{2}}$, where, I is a size K identity matrix. We used regularising informative priors setting $s = 1$ and $v = 4$.

These models are similar to those presented by (Jamil *et al.* 2013) and also evaluated by (Miller *et al.* 2019), described there as MLM1. Miller et al., found that having a fixed effect of traits improved model performance as in (Jamil *et al.* 2013), where there was an effect of traits on prevalence. Trait effects on prevalence should not be assumed, and much less work has demonstrated links between traits and commonness vs rarity (though see (Cornwell & Ackerly 2010). Whereas, much work has demonstrated associations between traits and gradients. Therefore in the spirit of parsimony we have retained the simpler model structure of excluding traits effects on prevalence.

Taxon models for target taxa and regions

To compare the predicted (trait-based) response of taxa to gradients in the target regions we fit simple logistic regressions of the form:

$$\Pr(Y_i = 1) = \text{logit}^{-1}(X_i \beta) \tag{2}$$

on per taxa, per region basis. Where X_i represented the same environmental gradients that were used in the Grampians (target) region trait-based model, but measured in the 18 target regions ≈25,000 plots. For each taxon in each region we then compared the estimated value of β (the taxon responses) to the predicted response attained by combining the traits of the target taxon with the parameters of equation 1 estimated from the target model.

Results

The magnitude of the (taxon-level) random effects can be seen to decline in the order: prevalence, Moisture Index, Ruggedness, Topographic Wetness Index and Total Nitrogen. The random effects are all correlated less than r = |0.5|.

Table S3.1: Random effects

	σ	ρ			
		B_0	MMI	TWI	R1k
B_0	1.2				
MMI	0.7	-0.2			
TWI	0.4	0	0.4		
R1k	0.5	-0.1	-0.4	-0.2	
TN	0.4	0.1	-0.4	-0.3	0.2

Fixed effects report the response of the hypothetical average taxon, with average traits, and then below, the trait-environment interactions.

Taxon coefficients for the GLMM trained in the Grampians illustrate considerable variation in response to Moisture Index, and rather less for Nitrogen (Fig. S3.1). Uncertainty in those coefficients is also greater for Moisture Index than Nitrogen. Prevalence (labelled intercept) varies widely among taxa.

Illustration of the predicted response surface across gradients of Moisture and Topographic Wetness is provided for four hypothetical eucalypts with contrasting combinations of SLA and seed mass (Fig. S3.2). It is clear these these hypothetical taxa would be differently distributed.

Table S3.2: Fixed effects

	μ	σ
B_0	-3.2	0.3
MMI	-0.3	0.2
TWI	0.1	0.1
R1k	0.4	0.1
TN	-0.1	0.1
MMI-SLA	-0.0	0.4
MMI-SM	0.2	0.3
MMI-MH	0.0	0.2
TWI-SLA	0.4	0.2
TWI-SM	-0.1	0.2
TWI-MH	0.0	0.1
R1k-SLA	0.5	0.3
R1k-SM	0.6	0.2
R1k-MH	-0.0	0.2
TN-SLA	-0.2	0.2
TN-SM	0.1	0.2
TN-MH	-0.1	0.1

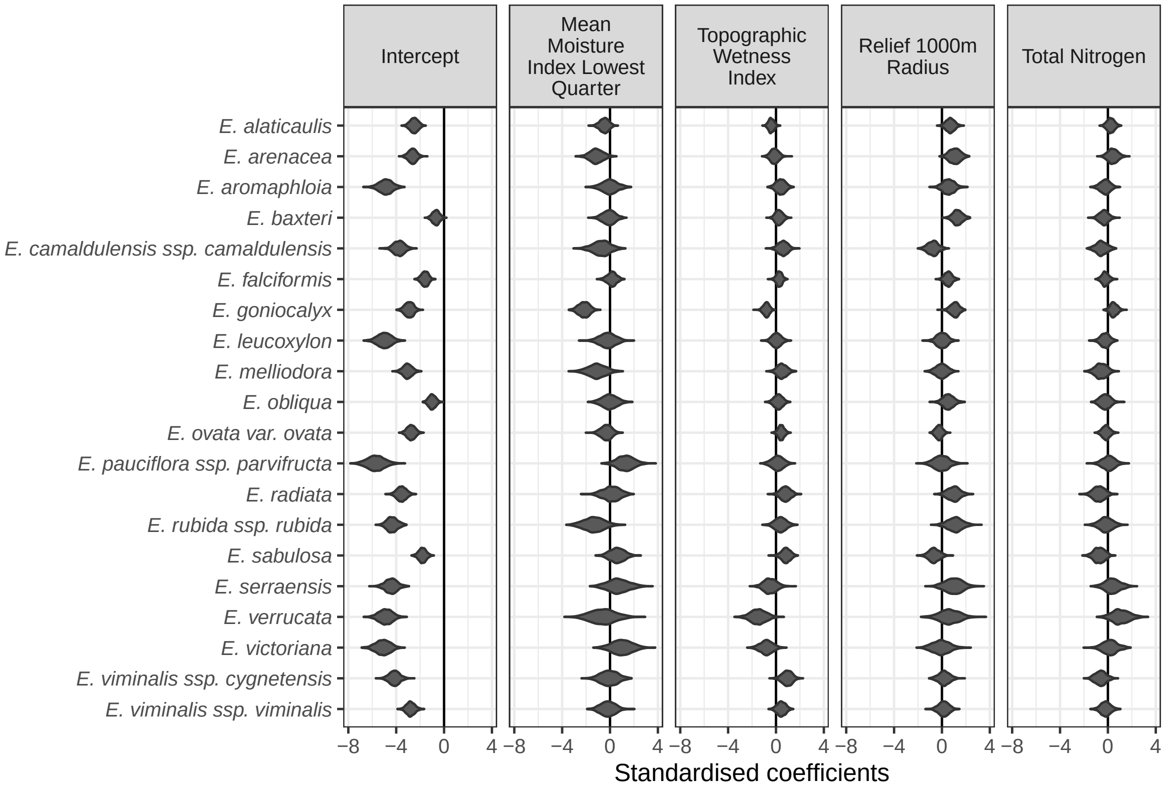

Figure S3.1: Estimates of taxon-specific model parameters. Each violin represents the uncertainty in the model intercept or environmental response coefficient for a taxon.

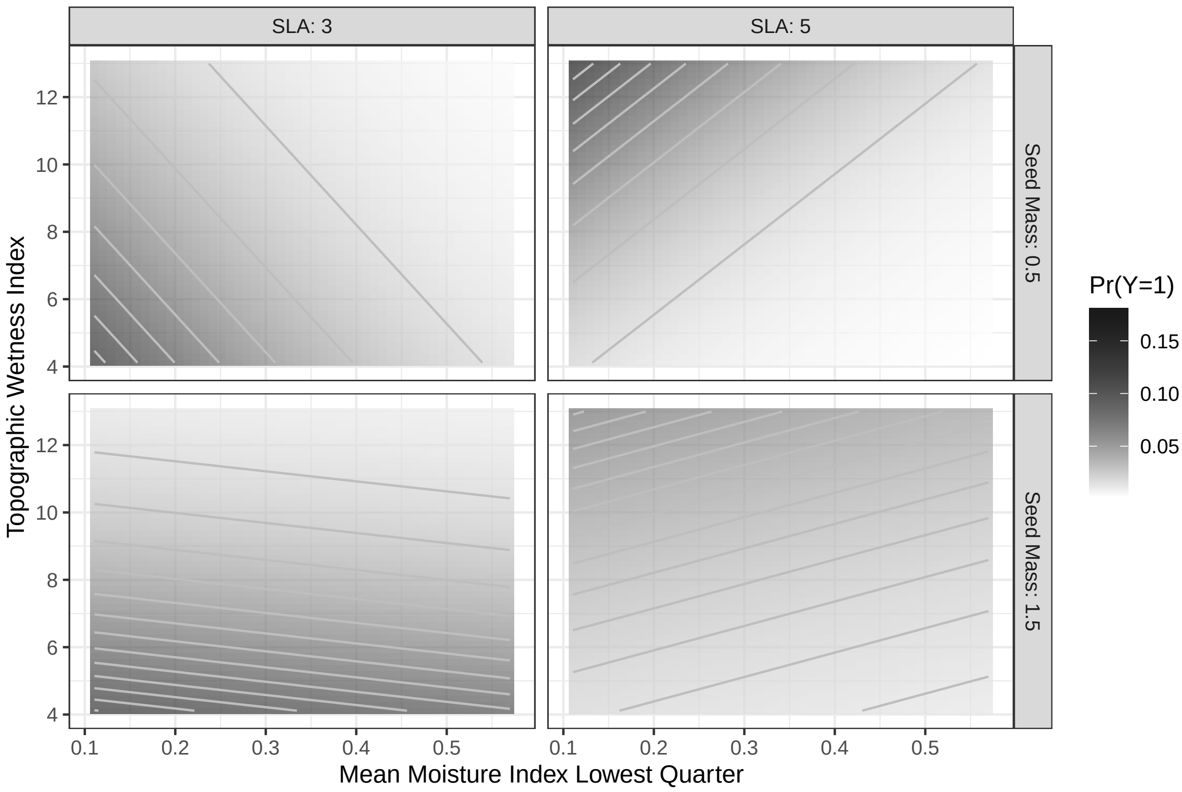

Figure S3.2: The predicted probability of occurrence in relation to two environmental gradients for four hypothetical eucalypt taxa with different combinations of two traits.

Partial responses of two exemplar taxa illustrate the fitted curves and occurrence data along the four environmental gradients (Fig. S3.3). Positive responses to Topographic Wetness can be seen despite no presences at high values, this can be due to a relative paucity of plots in such locations, and also that other gradients may covary with Topographic Wetness in such a way to overwhelm an apparent pattern with it.

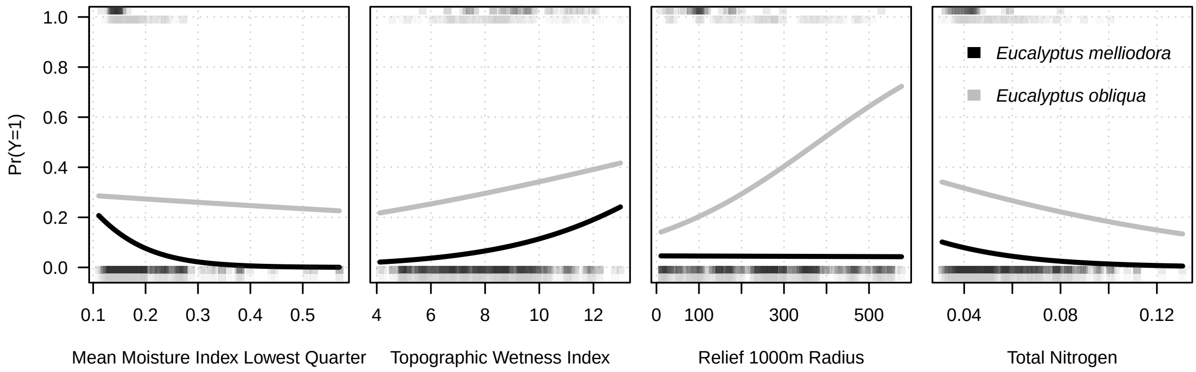

Figure S3.3: Expected partial response curves of two taxa from the Grampians trait-environment model.

Trait values in relation to modelled responses to one environmental gradient: topographic wetness.

In Fig. S3.4 the relationship of two traits (SLA and seed mass) to the responses to topographic wetness can be seen to be distributed along the 1:1 line between predicted coefficients and those estimated from taxon regressions. High and low trait values are found in opposite quadrants of the plots for the Grampians (top panels). Low SLA and high seed mass taxa are consistently found to have negative responses to Topographic Wetness, in the trait-SDM and taxon regressions.

In the Victorian Alps, we see most taxon responses captured well (Fig. S3.4, middle row). Two taxa lie to the top of the plot and left of the y-axis, indicating incorrectly predicted positive response, in keeping with their high SLA and middling seed mass.

But the taxon regression estimates indicate the taxa responding like taxa with lower SLA and/or heavier seeds. In both of the two target regions—Snowy Mountains and Victorian Alps—taxa have higher SLA and lighter seeds, with incorrectly positive responses predicted for several taxa. This suggests that trait ranges that extend well beyond the reference trait range might play some role in low predictive performance in testing ranges.

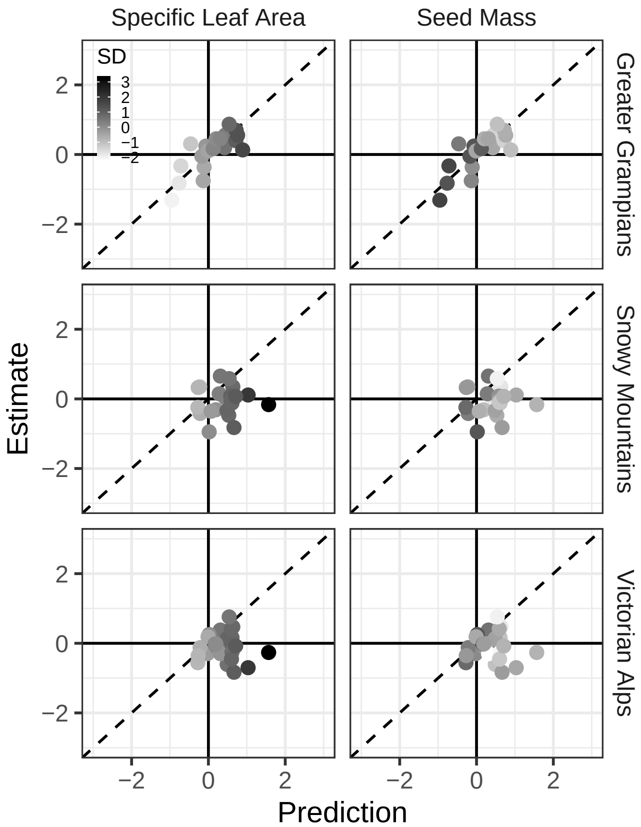

Figure S3.4: Trait-SDM predicted vs. single taxon model estimated response to topographic wetness gradient. Each point's black level is mapped to the taxon's median SLA or seed mass on a scale of standard deviations from the mean trait value of taxa in the Grampians.

References

Cornwell, W. K. & Ackerly, D. D. (2010). A link between plant traits and abundance: evidence from coastal California woody plants. *Journal of Ecology* 98, 814–821.

Jamil, T., Ozinga, W. A., Kleyer, M. & Braak, C. J. F. ter (2013). Selecting traits that explain species–environment relationships: a generalized linear mixed model approach. *Journal of Vegetation Science* 24, 988–1000.

Miller, J. E. D., Damschen, E. I. & Ives, A. R. (2019). Functional traits and community composition: A comparison among community-weighted means, weighted correlations, and multilevel models. *Methods in Ecology & Evolution* 10, 415–425.

Pollock, L. J., Kelly, L. T., Thomas, F. M., Soe, P., Morris, W. K., White, M. *et al.* (2018). Combining functional traits, the environment, and multiple surveys to understand semi-arid tree distributions. *Journal of Vegetation Science* 29, 967–977.

Pollock, L. J., Morris, W. K. & Vesk, P. A. (2012). The role of functional traits in species distributions revealed through a hierarchical model. *Ecography* 35, 716–725.