DR. LAURA J. POLLOCK (Orcid ID : 0000-0002-6004-4027)

Article type : Research article

Coordinating editor: Sándor Bartha

Title: Combining functional traits, the environment, and multiple surveys to understand semi-arid tree distributions

Authors: Laura J. Pollock^{1*}, Luke T. Kelly², Freya M. Thomas³, Paing Soe^{3,4}, William K. Morris³, Matt White⁵, and Peter A. Vesk³

Pollock, L.J. (laura.pollock@univ-grenoble-alpes.fr) * corresponding author

Kelly, L. T. (ltkelly@unimelb.edu.au)

Thomas, F. M. (freyaanthomas@gmail.com)

Paing Soe (ps2004@gmail.com)

Morris, W.K. (wkmor1@gmail.com)

White, M. (Matt.White@delwp.vic.gov.au)

Vesk, P.A. (pvesk@unimelb.edu.au)

¹ Univ. Grenoble Alpes, Laboratoire d'Écologie Alpine (LECA-CNRS), F-38000 Grenoble, France

²School of Ecosystem and Forest Sciences, The University of Melbourne, Victoria 3010, Australia

³School of BioSciences, ARC Centre of Excellence for Environmental Decisions, The University of Melbourne, Victoria 3010, Australia

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as [doi: 10.1111/jvs.12686](https://doi.org/10.1111/jvs.12686)

This article is protected by copyright. All rights reserved

⁴World Wildlife Fund-Myanmar, Dawei, Myanmar

⁵Arthur Rylah Inst. for Environmental Research, Dept of Environment, Land, Water and Planning 123 Brown Street, Heidelberg, Victoria 3084, Australia.

Type: Research Article

Printed journal page estimate : 6152 words, 2 tables, 4 figures

Abstract

Questions

Relationships between species, their functional traits and environmental gradients can now be more fully understood with trait-based multi-species distribution models (trait-SDMs). However, general patterns are yet to emerge from founding studies using these models, which are mostly case studies at a single scale. Here, we address the generality of trait-environment relations by asking whether these relationships hold for different sampling schemes, environmental variables, and species sets.

Methods

We focus on the drought and fire-resistant ‘mallee’ eucalypts of a semi-arid region of southeastern Australia, which are likely to face new climates and disturbance regimes under global change. We use hierarchical regression modeling to test how trait-environments change for two datasets representing an extensively collected, multi-purpose dataset and an intensively collected dataset stratified along environmental gradients.

Results

Three functional traits (specific leaf area, maximum height, and seed mass) explained a substantial portion of the occurrence of species along soil, water, and climatic gradients with the relationship between seed mass and soil type robust across all tests. Other trait-environment relationships changed depending on study design and species set with soil and substrate variables more important relative to climate (precipitation) for the intensively-sampled survey. Remotely-sensed variables were good surrogates for some field-based measures (soil type), but not others (landform: dune or swale). In particular, airborne soil radiometric data shows promise as a spatially-continuous substitute for soil texture.

Conclusions

Trait-SDMs are a powerful tool for quantifying ecological interactions, but generalisations will only be possible when sample design, scale, and environmental variables are carefully considered. We show that important ecological relationships can be diluted or missed entirely in broadscale trait-environment studies that rely on remotely sensed climate variables alone. Relationships that are robust to differences in study design, growth form and ecosystem (e.g. heavier seeds on sandy soil) are the most likely to reveal general ecological processes.

Keywords: Australia; *Eucalyptus*; hierarchical multi-species model; leaf-height-seed scheme; mallee; semi-arid ecosystem

Introduction

Plant functional traits capture fundamental physiological and morphological trade-offs faced by plants that have consequences for performance under different environmental conditions. Traits can help us learn why species occur where they do, and predict where species with similar traits may occur. The approach to understanding the relationships between functional traits and the environment has developed from characterizing trait distributions of communities at environmentally contrasted sites (Wright et al. 2001) to detailed quantitative analyses of how trait values vary within communities and along environmental gradients (Ackerly & Cornwell 2007). In many landscapes, the environmental backdrop is complex, with numerous overlain and interacting environmental gradients. The role of multiple environmental gradients can be captured with Species Distribution Models (SDMs), but the modeled relationships between species and the environment might arise from correlations with no underlying physiological basis—one of the strongest criticisms of SDMs. On the other hand, no single trait captures the autecology of a plant; treating a single trait as a response variable for a community ignores the influence of other traits on the occurrence of species within that community. The ability to model multiple functional traits interacting with multiple environmental gradients (trait-SDM) is a promising step forward in the quest for a more fundamental understanding of how species respond to environmental gradients and how the environment filters species into communities (Pollock et al. 2012; Jamil et al. 2013; Brown et al. 2014).

Early trait-SDM studies have established that traits do indeed explain a substantial portion of the variation in the way species respond to the environment (Pollock et al. 2012; Brown et al. 2014), that the predictive ability of these models is not typically diminished (and sometimes improved) compared with typical SDMs (Brown et al. 2014), and these models have some statistical advantages in checking model assumption and quantifying uncertainty compared with other approaches to the ‘fourth-corner’ problem (Jamil et al. 2013; Warton et al. 2014). However, previous work has mostly been on single datasets, with a single set of models and environmental variables. Many questions remain about whether trait-SDMs can detect functionally meaningful species-trait-environment relationships, especially as recent research suggests that the relationship between traits and growth rates might change from one region to another (Thomas & Vesk 2017). How trait-SDMs are affected by elements of study design such as spatial sampling design (Austin & Heyligers 1989) and choice of environmental covariates (Austin 2002), or species selection (Westoby 2002) remains under-studied and is the subject of this paper.

Ecologists know well that the type of sampling across space and the environment influences the inferences drawn from observational and experimental studies (Green 1979; Austin & Heyligers 1989; Jongman et al. 1995; Underwood 1997; Quinn & Keough 2002), and is a critical factor influencing ecological inference and predictive ability in SDM studies (Elith & Leathwick 2009). Therefore, the importance of sample design must extend to trait-environment modeling, and, here we investigate how combining surveys collected for different purposes offer insight into the generalizability of trait-environment relationships. We compare an extensively-sampled plot dataset collated for the purposes of regional- and state-wide vegetation classification (VBA: Liu et al. 2013) with an intensively-sampled study focused on fine-scale environmental responses within a broad vegetation type: mallee shrublands and woodlands (MBFP: Clarke et al. 2010). We document how trait-environment relationships change with sampling design and show how this change relates to different ecological processes operating at different scales.

Use of indirect environmental variables (*sensu* Austin 2002) to capture ecological processes may further obscure important trait-environment relations. This discrepancy of data type and ecological process is a major concern in the SDM field, and recent studies have shown that

edaphic variables (especially when measured on-site) can provide greater accuracy and predictive power than SDMs relying only on remotely-sensed variables (Luoto & Heikkinen 2008; Austin & Van Niel 2011; Bertrand et al. 2012). With increasingly diverse types of remotely-sensed data, and the attraction of linking them to compilations of species occurrence and trait datasets (e.g., (Wright et al. 2005; Moles et al. 2007; Moles et al. 2009; Kattge et al. 2011)), understanding the robustness of inferences based on remotely-sensed data is a critical concern. Here, we compare two remotely-sensed variables (remotely sensed Potassium and vertical height above streams) to field-measures of soil texture and landform. Species selection is also an important element of understanding trait and environment associations (Westoby 2002). As the species set under investigation is expanded, heterogeneous responses may be observed among species, which may be facing different challenges with appropriately distinct strategies. Plants growing in riparian areas within a semi-arid landscape may have quite different strategies from those in the wider landscape given the greater availability of soil moisture.

The overarching aim of this paper is to use trait-SDMs to understand the relationships between species, their traits and the environment for *Eucalyptus* species in semi-arid Australia. Our multi-scale case study data are from vegetation dominated by ‘mallee’ eucalypts with a short, multi-stemmed growth form that occur along relatively subtle soil, water and climatic gradients (White 2006). The mallee eucalypts occur across a wide area of Australia (the 200 - 500 mm rainfall zone) (Parsons & Rowan 1968), and are highly biologically diverse (Bradstock et al. 2002), but have been widely cleared for agriculture. We use the trait-SDM of Pollock *et al.* (2012), which combines traits from the leaf-height-seed scheme (Westoby 1998) with field-based and remotely-sensed environmental variables. This trait-SDM was originally tested in a more topographically-diverse area with steeper environmental gradients and *Eucalyptus* species with more varied traits. Each of three leaf-height-seed traits (specific leaf area (SLA), tree height and seed mass) were important in modulating species responses to different environmental gradients (Pollock et al. 2012). Testing the generality of these findings in a different system will provide a more robust understanding of trait-environment interactions.

We test hypotheses involving: the choice of covariates, the spatial sampling, and the species set. In the covariates hypothesis, we investigate whether the choice of covariates affects our ability to infer trait-environment relationships. Specifically, we test whether trait-environment relations found with field-measured edaphic variables can also be recovered with remotely-sensed variables. Second, we ask whether the intensity and spatial structure of sampling impacts the ability to infer trait-environment relations, and specifically test whether an extensively sampled multi-purpose dataset can recover the same effects as a survey with intensive sampling along environmental gradients. Third, we test whether trait-environment relationships are impacted by increasing the number and type of species considered by including eucalypts with different growth forms, and from different parts of the landscape (floodplain vs. sandsheets). We addressed each of these hypotheses by fitting models to each of two datasets that differed in one respect (spatial sampling, environmental covariates or species set).

Methods

Study area

The study area covers approximately 56,680 km² in southeastern Australia ranging 1.3° of latitude (Fig. 1). Large tracts of native vegetation are characterised by stands of ‘mallee’ vegetation (i.e. vegetation dominated by *Eucalyptus spp.* with a multi-stemmed growth form) (White 2006; Haslem et al. 2010). Native vegetation on floodplain areas is dominated by taller, single-stemmed eucalypts including *E. largiflorens* and *E. camaldulensis* (White 2006). Native vegetation has been widely cleared for agriculture outside protected areas with remnant vegetation largely restricted to less fertile soils and sand dunes (White 2006). The region is low in elevation (≤ 100 m above sea level) and is dominated by undulating dunes and swales bisected by the Murray-Darling river basin. The climate is semi-arid. Annual rainfall ranges from 230 mm in the north to 315 mm in the south. Summers are hot with maximum temperatures often exceeding 35°C and frosts are frequent in the winter. Higher temperatures and lower rainfall occur in the north and lower temperatures and higher rainfall in the south. Fire is a major driver of mallee vegetation structure, function and change (Bradstock et al. 2002), and wildfires exceeding 100 000 ha typically occur in the region every 10-20 years (Avitabile et al. 2013).

Eucalypt data

Eleven *Eucalyptus* species occur in the study area (see Table S1 in the Supplemental Information for a list of species). Eight species have a short (<10 m), multi-stemmed growth form (hereafter ‘mallee-form’). Three species have a tall (>10 m), single-stemmed growth form (hereafter ‘tree-form’). We sourced eucalypt occurrence records from two datasets: Victoria’s Biodiversity Atlas (VBA) (Department of Environment and Primary Industries) and The Mallee Fire and Biodiversity Project (MFBP) (Haslem et al. 2010). These two datasets cover roughly the same spatial extent, but represent different sampling strategies. VBA data are part of a larger dataset of 40,000 quadrats (or sample sites) collected over the past 40 years across the state of Victoria and neighboring areas for the purpose of vegetation classification. Here, we used the 3630 plots with eucalypt data in the study region (Fig. 1). Survey plots are typically 900 m² in size with size differences arising from local environmental variation. Plots are spread across the study area, but not systematically placed along particular environmental gradients. This serves as an extensive sampling design, covering a wide range of space and climate, but without intensive sampling within landscapes.

The MFBP dataset includes occurrence records of eight species of eucalypts from 598 sites in large nature reserves in Victoria and New South Wales (Haslem et al. 2010). We selected a subset of sites from the broader MFBP data that better matched the VBA data (i.e. we excluded data from South Australia). MFBP survey sites were clustered into landscapes, each a circular area 4 km in diameter (12.56 km²), selected to represent variation in fire regimes and stratified to sample dune-swale sequences (Clarke et al. 2010). In 2007, vegetation surveys were conducted once at each site in a quadrat 10 m x 50 m (500 m²) (Haslem et al. 2010). The MFBP data are restricted to vegetation on sandier soils, in conservation reserves, only contain mallee growth form species, but have detailed plot-based soil measurements and intensively sampling local landscape variation. By contrast, the VBA data includes vegetation on a wider range of soils, including those in agricultural areas, and both mallee and tree growth forms. These two datasets allow us to contrast trait-environment relationships across differing environmental ranges, for different growth forms, and to reveal whether an extensive plot dataset results in different inferences than intensive, stratified sampling.

Environmental covariates

We chose environmental covariates likely to be related to water, nutrients and light including field-measured (percent sand and dune landform), GIS-derived variables (winter precipitation, radioelement Potassium, and vertical height above streams). Radiometric potassium was used for broad-scale models as a surrogate for soil water availability, and correlates with various soil properties, especially soil texture (Pracilio et al. 2006; Read et al. 2018). For a detailed description of environmental covariates see Table S2. A comparison of ranges of the environmental covariates between the intensive dataset (MFBF) and the extensive dataset (VBA) is shown in Fig. 1. The links between the environmental covariates used in this study with hypothesized processes influencing species distributions are further explained in Fig. S1.

Trait data

We collected trait data on three plant functional traits that underpin the leaf-height-seed plant strategy scheme (Westoby 1998): specific leaf area (SLA), tree height and seed mass (see Table S1 for species trait values). Eucalypt leaves and fruits were collected in Murray Sunset NP, Wyperfeld NP and Little Desert NP in 2012-2013. Potential sites were located using field guides and recorded locations of species from both datasets, and we sampled 40 sites spread across the region. At each sample site, we collected trait data on three individuals of each species present. We quantified the trait values by taking the median value across all individuals of each species (for SLA and seed mass) and the maximum recorded value for height. We collected fully expanded leaves undamaged by insects from canopies exposed to direct sunlight. SLA and seed mass were measured in the laboratory following Pollock *et al.* 2012. Four species did not fruit during field surveys. For these species, we used University of Melbourne herbarium specimens instead (five individual fruit samples for each species). Maximum tree height was estimated using field samples (>5 for each species) and local field guides.

Description of models

We used a multi-species hierarchical GLM described in Pollock *et al.* (2012) to model how species respond to environmental gradients and how traits influence those responses. Trait-SDMs work by taking advantage of the fact that species respond differently to environmental

gradients (Vesk 2013). All species are simultaneously modelled (Gelfand et al. 2003; Ovaskainen & Soininen 2011; Pollock et al. 2012), with the expectation that individual species responses to environmental gradients vary (e.g. some species are found on sandy soils and others not). The main model level captures how each species varies with the environment, and is similar to a species distribution model (SDM) that predicts species occurrence as a function of environmental variables. Sub-models for the parameters for main predictor variables (e.g. soil texture) allow us to explicitly quantify whether the response of a species to an environmental predictor is a function of its traits (Pollock et al. 2012; Jamil et al. 2013; Vesk 2013; Brown et al. 2014). This model assumes linear responses of species to environmental variables, and traits are related to the slope of these responses. The model contains both variable slopes and variable intercepts, which allows parameters to vary for each species. We fitted the hierarchical model with REML estimation in R using the function ‘glmer’ in the package ‘lme4’ (Bates et al. 2015). All predictor variables and trait values were log- or square-root transformed as necessary and were centered and scaled by their standard deviations. We ran different versions of each model with and without random intercept terms (plots for VBA, and plots and landscape clusters for MFBF) to represent the spatial organization of each dataset as recommended in Jamil et al. (2013). Across both data sets, models with and without plots as a random effect were nearly indistinguishable, so we report only results from the models without a plot random effect. However, inclusion of a random intercept for landscape clusters in the MFBF dataset did improve model fit, so we report results for this version of MFBF models.

We fitted four models to datasets to address our three hypotheses (Table 1). Each hypothesis was tested comparing a pair of models fitted to datasets that differed in one respect at a time. Model 1 used the spatially-structured, intensively sampled MFBP dataset of mallee species (Table 2), which is the most comparable to that in Pollock *et al.* (2012) in that field-measured edaphic covariates represent local topography and soils, and climate is represented by remotely sensed data. Model 2 used the same species occurrence dataset, but replaced field-based percent sand and landform (Dune) with remotely sensed covariates: Radiometric Potassium and height above streams. Model 3 uses the extensive VBA dataset for mallee species (M3) and is extended to include tree species occurring on floodplains in Model 4 (M4) to explore the effect of species set in trait and environment relationships (Table 2). All

models had all three traits – SLA, tree height and seed mass - interacting with each environmental predictor. Our hypothesis tests constituted comparing the coefficients for effect of a trait on an environmental covariate with respect to the coefficient's sign, magnitude and uncertainty. We trialed models with and without polynomials terms, and found the results to be nearly identical (Fig. S2). We also compared the amount of variance described by the traits and environmental variables with the marginal and conditional R^2 statistic for random intercept models (Nakagawa & Schielzeth 2012) and modified for random slope models (Johnson 2014). For all parameter estimates 95% approximate credible intervals were calculated as 1.96 times the estimated standard error while for regression lines approximate 95% credible bands were calculated via simulation of parameter uncertainty given the estimated coefficient means and standard errors. See Appendix 1 for R code.

Results

All three traits modulated responses to some environmental variables in all models (Fig. 2), and the traits explained a substantial portion of the environmental responses for all (in terms of variance explained by the models). The addition of traits increased the marginal R^2 from 0.15 in Model 1 to 0.483 in Model 3 (Fig. S3). In the case of the models built on the intensively-sampled dataset with field-based edaphic covariates (M1), the strongest trait-environment interaction was maximum plant height and height above streams, whereby taller species increased in prevalence with increasing height above streams (M1; Fig. 2, circles). Taller species decreased in prevalence on dunes and sandy soils (Fig. 2). The most precise estimate in M1 was the interaction between seed mass and sandiness, whereby heavy-seeded species were more likely to be found on sandy soil (Fig. 2 top right panel) and dune crests. Species with high SLA (thinner, flimsier leaves) were more likely to be found on sandy soils with higher rainfall well above the surface stream network (Fig. 2).

The Covariates hypothesis

Prediction: if edaphic covariates measured in the field are correlated with covariates measured using remote sensing then species-trait interactions should be similar when modelled using either data

We tested whether remotely measured radiometric Potassium was a useful surrogate for field-based soil texture, and whether GIS-derived height above streams was a useful surrogate for

field-assessed position on the landscape (on a dune or not). Field-based and remotely-sensed variables were moderately correlated (Fig. S4) and we expect them to represent the same broad aspects of the environment. In the models, field-based and remotely-sensed edaphic variables produced most similar coefficients for SLA and seed mass, but not for height. We expected height would have the opposite interaction with remotely-sensed radiometric Potassium (higher Potassium generally means more clay and less sand) compared to field-measured % Sand and Dune. However, tall species responded negatively to all three, although the 95% credible interval overlaps zero for Height-Potassium (Fig. 2). While moderately correlated, height above streams was not a good surrogate for landform position (i.e. dune versus swale or flat) in the trait-SDMs, with contrasting trait-environment relationships observed (Fig. 2). For example, taller trees increased higher above streams, but decreased on dune tops (Fig. 2).

The Spatial Sampling hypothesis

Prediction: the intensity and spatial structure of sampling affects relationships between traits and environments

Spatial sampling structure had important effects on the sign, magnitude and uncertainty of trait-environment interactions. Those interactions involving maximum height were most affected, those with seed mass least affected and those with SLA intermediate. In two cases (SLA- height above streams, plant height-precipitation), the sign of the coefficient flipped from positive to negative or vice versa (Fig. 2). Trait-environment relationships inferred in the spatially-structured, intensive dataset (M2; Fig. 2, squares) tended to be dampened (smaller magnitude, less certain) in the extensive dataset (M3; Fig. 2, triangles) for edaphic variables (Potassium and height above streams, Fig. 2). For example, in both models, SLA negatively influenced species response to Potassium, but the response was more than twice as strong in the intensive sampling than extensive (Fig. 2, Fig. 3). The opposite was true for the climatic variable (precipitation), which became increasingly important in the extensive dataset covering a greater precipitation range, such as the relationship between SLA and precipitation (Fig. 2).

The Species Set hypothesis

Prediction: trait by environment relationships are affected by the number and type of species included in the model set

We found the interactions between SLA or seed mass and all environmental covariates were similar in direction and magnitude whether we considered only mallee eucalypts (M3, Fig. 2, triangles) or both mallee and floodplain eucalypts combined (M4, Fig. 2 inverted triangles). The point estimates of the parameters from each model lay within the corresponding parameter 95% credible bounds from the other model. When considering both growth forms, SLA was more negatively related to height above streams and was less positively related to precipitation. Similarly, seed mass became more strongly negatively correlated to Potassium and somewhat more negatively related to height above streams.

The differences between species subsets were more obvious when considering the effect of maximum height on species response to height above streams and precipitation. Taller trees were found on higher ground in the mallee model (M3), but the reverse is true for the model with all species—tall (single-stemmed) species occur closer to rivers (Figs. 2 & 4). Similarly, taller species are found at higher rainfall for the mallee models, but at lower rainfall in the model including both mallee and floodplain eucalypts.

The variable height above streams was particularly prone to changing responses across the different models (Fig. 4). There was an overall shift from tall species with high SLA being found higher above streams at the small scale with only mallees to tall, high SLA species being closer to streams at a large scale when all species are included (Fig. 4).

Discussion

Trait-SDMs are a powerful tool for quantifying ecological interactions, but robust inferences will only be possible when important aspects of sampling design and data are considered. Our results show that trait-environment relationships can change with spatial sampling design, choice of variables, and species set. Sampling a broader spatial and environmental extent less intensively may weaken some trait-environment interactions (especially for edaphic variables) and strengthen others such as climate (see Fig. 2). Different sets of plant species may have different trait-environment interactions, and different underlying relationships between the traits and physiological processes such as growth rates (Thomas & Vesk 2017).

This means a solid understanding of environmental drivers is essential for interpreting trait-based models. Trait-environment relationships that transcend study design and scale have greater power to explain how plants interact with their environment through traits. We discuss these results more fully below.

1. Remotely-sensed variables can be effective surrogates for field-based measures

Recent SDM studies suggest that site measurements result in more accurate predictions than models using broad scale variables (Austin & Van Niel 2011). This should be an important consideration for ecologists using SDMs, as predictive accuracy will have to be balanced against time and financial costs of collecting field-based variables. In the present study, we compared airborne radiometric data to field-based soil and landform variables. Radiometric data are promising for modelling species distributions because they offer spatially continuous coverage and thus a potential proxy for functional soil properties, which are crucial drivers of plant composition but are often missing from SDMs. Radiometric Potassium is correlated to soil texture (Pracilio et al. 2006; Read et al. 2018), and has been a powerful predictor variable for vegetation studies in semi-arid southeastern Australia (Read et al. 2008). Here, we found that radiometric data offers a useful surrogate for field-based measures of soil texture and landform. The main trait-environment interaction that was not apparent when using radiometric data, but was apparent with field-based model covariates, was the relationship with maximum height and soil type. We believe this difference for maximum height might result from the inability of Radiometric Potassium to distinguish soil nutrients and salinity in very low-lying, clayey basins where salts accumulate in this system. Shorter trees might replace taller trees that would normally inhabit clayey soils in situations where salinity is destructively high. Also, some narrow landscape features in the dune and swale system are likely to be missed with radiometric sensing.

The other comparison of remotely-sensed and field-based variables was between field-based landform classification and height above streams. We found height above streams to be an ineffective surrogate for landform (dune, flat, swale). However, we found height above streams to be a useful surrogate for another fine-scale gradient of water availability. For this reason, we retained it in the intensively-sampled dataset (M1) and it was the strongest effect of maximum plant height-height above streams. We interpret height above streams to

represent the vertical distance above groundwater and clay subsoils or, in other words, the depth of the sand-sheet within which soil moisture may be stored. Therefore, in an arid and streamless context, it probably represents greater water availability and enables growth of taller trees with deep root systems and higher SLA. At a larger spatial extent, especially when single-stemmed trees are included, the floodplain systems overwhelm this sand-sheet soil moisture, with tall, high SLA trees being closer to streams.

2. Trait interactions with local edaphic properties are better resolved with intentionally stratified, intensive sampling, broad-scale climate with extensive sampling

In the extensively sampled dataset, precipitation was more important in the case of SLA and height, while Potassium and height above streams tended to be less important. The broad-scale relationships that height and SLA both increase with precipitation are consistent with global studies (Wright et al. 2004; Moles et al. 2009), and within Australia for eucalypts in a more diverse, mesic region (Pollock et al. 2012). Additionally, other leaf traits, such as secondary compounds, are also known to influence moisture relations in mallee eucalypt species (Merchant et al. 2007) but it remains unclear how these traits relate to the more commonly measured traits that reflect general ecological strategies (Cornelissen et al. 2003).

Precipitation was more important in the extensively collected multi-purpose dataset, likely because some of the signal of edaphic properties we know to be important in this system were diluted or lost as the fine scale gradients were not resolved. The stratified sampling along known environmental gradients, concentrated in local areas, means that there is a better chance that the environmental values represent the actual gradient of importance. If sites are spread more sparsely, there is a chance that other factors (e.g. dispersal limitation) obscure this signal.

3. Some aspects of trait-environment for mallee growth form are consistent with those of floodplain trees, others reveal contrasting mallee ecology

Trait-environment interactions with SLA and seed mass were generally consistent between the mallee eucalypt species set compared with mallee and floodplain trees combined. However, trait-environment interactions with height differed between species sets. Taller mallee eucalypts were more likely to occur on higher ground (because of the deeper sand

440 sheets on dunes), but the tallest single-stemmed trees were more likely to occur on river
441 floodplains. Floodplain trees likely had great leverage on the relationships involving height,
442 because they have relatively extreme heights compared with mallee species. Yet such an
443 effect was not observed with SLA, despite trees having higher SLA than the mallee
444 eucalypts. Although our species set is restricted to the tree composition of the region, which
445 happens to consist of one genus, and the mallee growth form is peculiar to Australia, we
446 believe the results could reveal trait-environment relations found more broadly in semi-arid
447 landscapes where trees are near their physiological limits. Given the trees are relatively
448 closely related, divergences in traits are likely to be habitat-driven rather than legacy traits
449 that differ because they are from very different lineages.

451 *4. Seed mass and soil texture relationship robust to study design and region*

452 The mallee eucalypts that dominate this landscape are functionally similar at first glance—
453 they are short trees with multiple stems and the ability to vigorously re-sprout after fire.
454 Nevertheless, slight differences in their functional traits strongly influence the distribution of
455 these species in the landscape. Soil and landform are especially important at a small scale.
456 Short species with high SLA (flimsy leaves) and heavy seeds tend to be found on sandy soils
457 and dune crests. Shorter eucalypts typically have heavier seeds (Murray & Gill 2001). The
458 link between heavy seeds and sandy soils was consistent across scales, across species sets,
459 and using both field-based measures and a remotely-sensed variable. This relationship is also
460 consistent with results from a more topographically-complex and mesic environment
461 (Pollock et al. 2012). The importance of soil texture on mallee species occurrence has long
462 been known (Parsons & Rowan 1968), but our detailed modelling helps to explain why.
463 Seedling survival is generally greater for species with larger seeds (Leishman et al. 2000;
464 Moles & Westoby 2004), which probably translates to heavy seeds offering more advantages
465 in infertile sandy soils. In more mesic environments, smaller-seeded species likely
466 outcompete larger-seeded species because of their increased dispersal ability (Cremer 1977)
467 and increased fecundity. We hypothesise the seed mass-soil texture relationship exists due to
468 a tolerance-fecundity trade-off (Muller-Landau 2010), and heavy seeds might be the key to
469 survival and persistence in these harsh, sandy environments at the extreme of where trees can
470 grow.

Conclusions

Our results confirm that trait-based SDMs are a powerful approach for a refined understanding of how the environment filters species through their traits. This is true even along the relatively subdued environmental gradients of our case study region. We showed that heavier seeded species prefer sandier soils in semi-arid mallee vegetation. The link between seed mass and soil texture is consistent with modelling from more topographically diverse parts of Australia, and we hypothesise that this relationship might extend to other regions and plant groups. Our modelling reinforces the important role that edaphic properties play as fundamental drivers of vegetation, as well as showing that airborne radiometric data can provide spatially continuous coverage of these crucial, but often overlooked, soil properties. Finally, the way data are collected can strongly influence the outcome of trait-environment relationships. We recommend that future trait-based work consider the effect of taxonomic, functional and spatial design aspects, so that we can arrive at more general conclusions about how plants respond to their environments and can predict distributions of new species in new environments from their traits.

Acknowledgements

We acknowledge funding from the Bjarne K. Dahl Trust, the Australian Research Council Centre of Excellence in Environmental Decisions, and the National Environmental Research Program (NERP). We thank Amy Whitehead and José Lahoz-Monfort for help in the field and Gillian Brown and the University of Melbourne Herbarium for providing samples for trait measurements. Many thanks to the Mallee Fire and Biodiversity Project for providing eucalypt occurrence data.

References

- Ackerly, D.D., & Cornwell, W.K. 2007. A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecology Letters* 10: 135–145.
- Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* 157: 101–118.
- Austin, M.P., & Heyligers, P.C. 1989. Vegetation survey design for conservation:

- 504 GRADSECT sampling of forests in north-eastern NSW. *Biological Conservation* 50:
505 13–32.
- 506 Austin, M.P., & Van Niel, K.P. 2011. Improving species distribution models for climate
507 change studies: variable selection and scale. *Journal of Biogeography* 38: 1–8.
- 508 Avitabile, S.C., Callister, K.E., Kelly, L.T., Haslem, A., Fraser, L., Nimmo, D.G., Watson,
509 S.J., Kenny, S.A., Taylor, R.S., Spence-Bailey, L.M., Bennett, A.F., & Clarke, M.F.
510 2013. Systematic fire mapping is critical for fire ecology, planning and management: A
511 case study in the semi-arid Murray Mallee, south-eastern Australia. *Landscape and*
512 *Urban Planning* 117: 81–91.
- 513 Bates, D., Maechler, M., Bolker, B., & Walker, S. 2015. *lme4: Linear mixed-effects models*
514 *using Eigen and S4. R package version 1.1-9.*
- 515 Bertrand, R., Perez, V., & Gégout, J.-C. 2012. Disregarding the edaphic dimension in species
516 distribution models leads to the omission of crucial spatial information under climate
517 change: the case of *Quercus pubescens* in France. *Global Change Biology* 18: 2648–
518 2660.
- 519 Bradstock, R.A., Cohn, J.S., Williams, J.E., & Gill, A.M. 2002. Fire regimes and biodiversity
520 in semi-arid mallee ecosystems. In *Flammable Australia: the fire regimes and*
521 *biodiversity of a continent*, pp. 238–258.
- 522 Brown, A.M., Warton, D.I., Andrew, N.R., Binns, M., Cassis, G., & Gibb, H. 2014. The
523 fourth-corner solution—using predictive models to understand how species traits interact
524 with the environment. *Methods in Ecology and Evolution* 5: 344–352.
- 525 Clarke, M.F., Avitabile, S.C., Brown, L., Callister, K.E., Haslem, A., Holland, G.J., Kelly,
526 L.T., Kenny, S.A., Nimmo, D.G., Spence-Bailey, L.M., Taylor, R.S., Watson, S.J., &
527 Bennett, A.F. 2010. Ageing mallee eucalypt vegetation after fire: insights for
528 successional trajectories in semi-arid mallee ecosystems. *Australian Journal of Botany*
529 58: 363–372.
- 530 Cornelissen, J.H.C., Cerabolini, B., Castro-Díez, P., Villar-Salvador, P., Montserrat-Martí,
531 G., Puyravaud, J.P., Maestro, M., Werger, M.J.A., & Aerts, R. 2003. Functional traits
532 of woody plants: correspondence of species rankings between field adults and
533 laboratory-grown seedlings? *Journal of Vegetation Science* 14: 311–322.
- 534 Cremer, K.W. 1977. Distance of seed dispersal in eucalypts estimated from seed weights.
535 *Australian Forest Research* 7: 225–228.

- 536 Elith, J., & Leathwick, J.R. 2009. Species Distribution Models: Ecological Explanation and
537 Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and*
538 *Systematics* 40: 677–697.
- 539 Gelfand, A.E., Silander Jr., J.A., Wu, S., Latimer, A., Lewis, P.O., Rebelo, A.G., & Holder,
540 M.T. 2003. Explaining species distribution patterns through hierarchical modeling.
541 *Bayesian Analysis* 1: 1–35.
- 542 Green, R.H. 1979. *Sampling Design and Statistical Methods for Environmental Biologists*.
543 John Wiley & Sons.
- 544 Haslem, A., Callister, K.E., Avitabile, S.C., Griffioen, P.A., Kelly, L.T., Nimmo, D.G.,
545 Spence-Bailey, L.M., Taylor, R.S., Watson, S.J., Brown, L., & others. 2010. A
546 framework for mapping vegetation over broad spatial extents: a technique to aid land
547 management across jurisdictional boundaries. *Landscape and urban planning* 97: 296–
548 305.
- 549 Jamil, T., Ozinga, W.A., Kleyer, M., & ter Braak, C.J.F. 2013. Selecting traits that explain
550 species–environment relationships: a generalized linear mixed model approach. *Journal*
551 *of Vegetation Science* 24: 988–1000.
- 552 Johnson, P.C.D. 2014. Extension of Nakagawa & Schielzeth’s R2GLMM to random slopes
553 models. *Methods in Ecology and Evolution* 5: 944–946.
- 554 Jongman, R.H.G., Ter Braak, C.J.F., & Van Tongeren, O.F.R. 1995. *Data analysis in*
555 *community and landscape ecology*. Cambridge Univ. Press.
- 556 Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönsch, G., Garnier, E.,
557 Westoby, M., Reich, P.B., Wright, I.J., Cornelissen, J.H.C., Violle, C., Harrison, S.P.,
558 Bodegom, V., M, P., Reichstein, M., Enquist, B.J., Soudzilovskaia, N.A., Ackerly,
559 D.D., Anand, M., Atkin, O., Bahn, M., Baker, T.R., Baldocchi, D., Bekker, R., Blanco,
560 C.C., Blonder, B., Bond, W.J., Bradstock, R., Bunker, D.E., Casanoves, F.,
561 Cavender-Bares, J., Chambers, J.Q., Iii, C., S, F., Chave, J., Coomes, D., Cornwell,
562 W.K., Craine, J.M., Dobrin, B.H., Duarte, L., Durka, W., Elser, J., Esser, G., Estiarte,
563 M., Fagan, W.F., Fang, J., Fernández- Méndez, F., Fidelis, A., Finegan, B., Flores, O.,
564 Ford, H., Frank, D., Freschet, G.T., Fyllas, N.M., Gallagher, R.V., Green, W.A.,
565 Gutierrez, A.G., Hickler, T., Higgins, S.I., Hodgson, J.G., Jalili, A., Jansen, S., Joly,
566 C.A., Kerkhoff, A.J., Kirkup, D., Kitajima, K., Kleyer, M., Klotz, S., Knops, J.M.H.,
567 Kramer, K., Kühn, I., Kurokawa, H., Laughlin, D., Lee, T.D., Leishman, M., Lens, F.,

- Lenz, T., Lewis, S.L., Lloyd, J., Llusià, J., Louault, F., Ma, S., Mahecha, M.D.,
Manning, P., Massad, T., Medlyn, B.E., Messier, J., Moles, A.T., Müller, S.C.,
Nadrowski, K., Naeem, S., Niinemets, Ü., Nöllert, S., Nüske, A., Ogaya, R., Oleksyn,
J., Onipchenko, V.G., Onoda, Y., Ordoñez, J., Overbeck, G., Ozinga, W.A., Patiño, S.,
Paula, S., Pausas, J.G., Peñuelas, J., Phillips, O.L., Pillar, V., Poorter, H., Poorter, L.,
Poschlod, P., Prinzing, A., Proulx, R., Rammig, A., Reinsch, S., Reu, B., Sack, L.,
Salgado- Negret, B., Sardans, J., Shiodera, S., Shipley, B., Siefert, A., Sosinski, E.,
Soussana, J.-F., Swaine, E., Swenson, N., Thompson, K., Thornton, P., Waldram, M.,
Weiher, E., White, M., White, S., Wright, S.J., Yguel, B., Zaehle, S., Zanne, A.E., &
Wirth, C. 2011. TRY – a global database of plant traits. *Global Change Biology* 17:
2905–2935.
- Leishman, M., Wright, I.J., Moles, A.T., & Westoby, M. 2000. The evolutionary ecology of
seed size. In Fenner, M. (ed.), *Seeds: the ecology of regeneration in plant communities*,
2nd Edition, pp. 31–57. CABI Publishing, Wallingford, UK.
- Liu, C., White, M., Newell, G., & Griffioen, P. 2013. Species distribution modelling for
conservation planning in Victoria, Australia. *Ecological Modelling* 249: 68–74.
- Luoto, M., & Heikkinen, R.K. 2008. Disregarding topographical heterogeneity biases species
turnover assessments based on bioclimatic models. *Global Change Biology* 14: 483–
494.
- Merchant, A., Callister, A., Arndt, S., Tausz, M., & Adams, M. 2007. Contrasting
physiological responses of six Eucalyptus species to water deficit. *Annals of Botany*
100: 1507–1515.
- Moles, A.T., Ackerly, D.D., Tweddle, J.C., Dickie, J.B., Smith, R., Leishman, M.R.,
Mayfield, M.M., Pitman, A., Wood, J.T., & Westoby, M. 2007. Global patterns in seed
size. *Global Ecology and Biogeography* 16: 109–116.
- Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W., Zanne, A.E., Pitman,
A., Hemmings, F.A., & Leishman, M.R. 2009. Global patterns in plant height. *Journal*
of Ecology 97: 923–932.
- Moles, A.T., & Westoby, M. 2004. Seedling survival and seed size: a synthesis of the
literature. *Journal of Ecology* 92: 372–383.
- Muller-Landau, H.C. 2010. The tolerance–fecundity trade-off and the maintenance of
diversity in seed size. *Proceedings of the National Academy of Sciences* 107: 4242–

4247.

- Murray, B.R., & Gill, A.M. 2001. A comparative study of interspecific variation in fruit size among Australian eucalypts. *Ecography* 24: 651–658.
- Nakagawa, S., & Schielzeth, H. 2012. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution* 4: 133–142.
- Ovaskainen, O., & Soininen, J. 2011. Making more out of sparse data: hierarchical modeling of species communities. *Ecology* 92: 289–295.
- Parsons, R.F., & Rowan, J.N. 1968. Edaphic range and cohabitation of some mallee eucalypts in south-eastern Australia. *Australian Journal of Botany* 16: 109–116.
- 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.
- Pracilio, G., Adams, M.L., Smettem, K.R.J., & Harper, R.J. 2006. Determination of Spatial Distribution Patterns of Clay and Plant Available Potassium Contents in Surface Soils at the Farm Scale using High Resolution Gamma Ray Spectrometry. *Plant and Soil* 282: 67–82.
- Quinn, G.P., & Keough, M.J. 2002. *Experimental Design and Data Analysis for Biologists*. Cambridge University Press.
- Read, C.F., Duncan, D.H., Ho, C.Y.C., White, M., & Vesk, P.A. 2018. Useful surrogates of soil texture for plant ecologists from airborne gamma-ray detection. *Ecology and Evolution* 8: 1974–1983.
- Read, C.F., Duncan, D.H., Vesk, P.A., & Elith, J. 2008. Biological soil crust distribution is related to patterns of fragmentation and landuse in a dryland agricultural landscape of southern Australia. *Landscape Ecology* 23: 1093–1105.
- Thomas, F.M., & Vesk, P.A. 2017. Are trait-growth models transferable? Predicting multi-species growth trajectories between ecosystems using plant functional traits. *PLOS ONE* 12: e0176959.
- Underwood, A.J. 1997. *Experiments in Ecology: Their Logical Design and Interpretation Using Analysis of Variance*. Cambridge University Press.
- Vesk, P.A. 2013. How traits determine species responses to environmental gradients. *Journal of Vegetation Science* 24: 977–978.
- Warton, D.I., Foster, S.D., Déath, G., Stoklosa, J., & Dunstan, P.K. 2014. Model-based

- thinking for community ecology. *Plant Ecology*. doi: 10.1007/s11258-014-0366-3
- Westoby, M. 1998. A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil* 199: 213–227.
- Westoby, M. 2002. Choosing species to study. *Trends in ecology and evolution* 17: 587.
- White, M.D. 2006. The Mallee vegetation of north western Victoria. *Proceedings of the Royal Society of Victoria* 118: 229–243.
- Wright, I.J., Reich, P.B., Cornelissen, J.H.C., Falster, D.S., Groom, P.K., Hikosaka, K., Lee, W., Lusk, C.H., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Warton, D.I., & Westoby, M. 2005. Modulation of leaf economic traits and trait relationships by climate. *Global Ecology and Biogeography* 14: 411–421.
- Wright, I.J., Reich, P.B., & Westoby, M. 2001. Strategy shifts in leaf physiology, structure and nutrient content between species of high- and low-rainfall and high- and low-nutrient habitats. *Functional Ecology* 15: 423–434.
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, U., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J., & Villar, R. 2004. The worldwide leaf economics spectrum. *Nature* 428: 821–827.

List of Supporting Information

Table S1. Description of *Eucalyptus* species of the mallee region used in this study and trait values.

Table S2. Description of spatial and field-based environmental variables.

Figure S1. Influence diagram of hypothesized relationships between traits environmental variables, and fundamental processes underlying species distributions.

Figure S2. Interaction coefficients for each environmental variable and trait-environment combination for Model 1 constructed with polynomial terms (for comparison with Fig. 2 in the main text).

Figure S3. The effect of traits by environment interactions on marginal R^2 .

Figure S4. Relationships between field-based variables (% Sand and Landform) and their remotely-sensed surrogates (radiometric Potassium and Distance above streams) in the “Covariates Hypothesis”.

Appendix 1 R code

Table 1. Hypothesis tests and corresponding model comparisons (see Table 2 for model descriptions

Hypothesis	Comparison
H1, Covariates. If field-based (more proximal) and remotely-sensed (more distal) edaphic covariates are correlated, then M1 should have similar model parameters to M2.	M1 versus M2
H2, Spatial Sampling. If trait-environment relationships are insensitive to the location and breadth of sampling, then M2 should have similar model parameters to M3.	M2 versus M3
H3, Species Set. If trait-environment relationships are general for different species sets, then M3 and M4 should have similar parameters	M3 versus M4

Table 2. Description of models, variables* and which hypotheses they are used to test.

Model	Hypotheses	Species set	Edaphic covariates	Dataset
M1-Field-based variables	H1 - Covariates	‘Mallee’	Field-assessed Sand content + Dune	MFBP (intensive sampling)
M2-intensive	H1, H2 –	‘Mallee’	Remote sensed Potassium	MFBP

sampling	Covariates/Spatial Sampling			(intensive sampling)
M3-extensive sampling	H2, H3 – Spatial Sampling/Species Set	‘Mallee’	Remote sensed Potassium	VBA (extensive sampling)
M4-extensive sampling + floodplain trees	H3 – Species Set	‘Mallee’ and ‘floodplain tree’	Remote sensed Potassium	VBA (extensive sampling)

*all models include vertical height above streams and precipitation

Figures

Fig. 1. Map of the study area in northwestern Victoria, Australia (34°53'S 141°34'E) showing plot locations for the smaller scale ecological survey (MFBP) and the larger-scale survey sites from Victoria's Biodiversity Atlas (VBA). Several large conservation reserves are shown in grey: Murray-Sunset National Park (6330 km²), Hattah Kulkyne NP (480 km²), Wyperfeld NP (3568 km²), Big Desert (1423 km²) and Little Desert NP (1326 km²). Bargraphs are the range of trait and environment values (only those variables used at both scales) for the MFBP dataset (red bars) compared to the VBA dataset. See Tables S1 & S2 for more information on traits and variables.

Model 1 = ● Model 2 = ■ Model 3 = ▲ Model 4 = ▼

Response to vertical
height above streams

um height

Fig. 2. Interaction coefficients for each trait-environment combination for all four models: Intensively-sample, field (Model 1) and remotely-sensed variables (Model 2), extensively sampled, mallee-only dataset (Model 3) and combined mallee and tree set (Model 4). Interaction coefficients describe how traits modulate species occupancy relative to environmental variables (given other traits and environmental variables are held at their means). Positive coefficients mean that larger trait values increase the probability of species occurrence along an increasing environmental gradient. Negative values indicate a decrease in occurrence. Bars are credible intervals around parameter estimates, and we consider effects significant if the 95% credible intervals do not overlap zero. Each of the coefficients can be visualized, as examples show for cross species means and 95% credible intervals in the right-hand panels. The top panel shows the positive coefficient for sand content and seed mass, in which species responses to sand content is explained by seed mass for mallee eucalypt species. The lower panel depicts how taller species have more positive occurrence responses as height above streams increases. Values on the y-axis indicate the change in expected probabilities of species occurrence given a change in sand or change in precipitation. For

example, as seed mass increases across species, we expect an increased probability of that species occurring given an increase in percentage sand.

Fig. 3. The dilution of effect strength in the large-scale dataset illustrated by the SLA and Potassium interaction in the intensively-sampled Model 2 (left) and extensive Model 3 (right) datasets with violin plots depicting probability densities. Values on the y-axis indicate the change in expected probabilities of species occurrence given a change in Potassium. Species fall in similar relative locations in each graph, but the interactive effect weakens when a wider environment is represented.

Fig. 4. How trait relationships with vertical height above streams vary between models (see Fig. 2, Table 2). The effect of species traits in influencing species occurrence along a gradient of height above streams can weaken or, in fact, flip depending on the spatial extent and growth form extent of the dataset. For example, maximum plant height increases the probability of occurrence with increasing height above streams (Model 1). This effect weakens in the extensively sampled dataset (Model 3), and is reversed when all tree species are included (Model 4).

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Pollock, LJ;Kelly, LT;Thomas, FM;Soe, P;Morris, WK;White, M;Vesk, PA

Title:

Combining functional traits, the environment and multiple surveys to understand semi-arid tree distributions

Date:

2018-11-01

Citation:

Pollock, L. J., Kelly, L. T., Thomas, F. M., Soe, P., Morris, W. K., White, M. & Vesk, P. A. (2018). Combining functional traits, the environment and multiple surveys to understand semi-arid tree distributions. JOURNAL OF VEGETATION SCIENCE, 29 (6), pp.967-977. <https://doi.org/10.1111/jvs.12686>.

Persistent Link:

<http://hdl.handle.net/11343/284875>