

# Fine-scale hydrological niche differentiation through the lens of multi-species co-occurrence models

Andrew D. Letten<sup>1,\*</sup>, David A. Keith<sup>1,2,3</sup>, Mark G. Tozer<sup>2</sup> and Francis K.C. Hui<sup>4</sup>

<sup>1</sup>Centre for Ecosystem Science, School of Biological, Earth and Environmental Sciences, UNSW, Sydney NSW 2052, Australia; <sup>2</sup>NSW Office of Environment and Heritage, Hurstville NSW 2220, Australia; <sup>3</sup>Long Term Ecological Research Network, Terrestrial Ecosystem Research Network, Fenner School of the Environment, Australian National University, Canberra ACT 2601, Australia; and <sup>4</sup>School of Mathematics and Statistics, UNSW, Sydney NSW 2052, Australia

## Summary

**1.** Theory suggests spatial heterogeneity can facilitate species co-occurrence at fine scales, but environmental data are rarely collected at sufficiently high resolution to test this empirically. Whilst there is emerging evidence that subtle variation in soil hydrology represents a fundamental fine-scale niche axis within plant communities, this is largely derived from studies of soil hydrology in isolation from other environmental factors.

**2.** We assessed the comparative importance of fine-scale hydrological niche differentiation for species co-occurrence using a high-resolution study of soil hydrology and other edaphic variables, coupled with a long-term (24 years) data set of herbaceous plant plots in a heathland community in south-east Australia.

**3.** For the analysis, we employed novel latent variable models (LVMs), which offer an explicit, model-based approach to partitioning out the different drivers of species co-occurrence patterns. Whilst the regression component of an LVM models the species-specific environmental responses, the latent variable component can be used to identify residual patterns of co-occurrence, which may be attributable to unmeasured factors and/or biotic interactions.

**4.** Relative to a host of plant resources, non-resource factors and ‘unmeasured’ latent variables, soil hydrology emerged as the best predictor of negative co-occurrences within the community, with the dominant species exhibiting strongly differentiated responses across a comparatively narrow moisture gradient. Nevertheless, strong species-specific responses to environmental variability only emerged at scales greater than those at which plants may be expected to compete for resources, throwing doubt on the direct role of spatial heterogeneity as a mechanism for local-scale coexistence.

**5. Synthesis.** This study confirms the vital role of hydrological niches for the maintenance of within-community plant diversity, but also highlights the need for more rigorous analysis of scale dependencies to better understand the underlying coexistence mechanisms at play. In addition, it illustrates the inferential gains made possible with model-based approaches to the analysis of species co-occurrence. R code illustrating model fitting and inference is provided as a supplement.

**Key-words:** coexistence, community assembly, determinants of plant community diversity and structure, hydrological niche, latent variables, multi-species models

## Introduction

Spatial environmental heterogeneity is widely recognized to play a fundamental role in driving broad-scale patterns of diversity and compositional turnover. The underlying mechanisms are diverse, but include wider niche space, more refuges from adverse conditions, and stronger selection for evolutionary diversification (reviewed in Stein, Gerstner & Kreft 2014). At finer, within-community scales, however, the

extent to which heterogeneity explains observed patterns of species co-occurrence is less clear. In spite of a wealth of theory suggesting that ‘variation-dependent’ coexistence mechanisms may be pervasive (Chesson 2000; Amarasekare 2003; Snyder & Chesson 2004), the empirical evidence is surprisingly sparse (but see Fridley et al. 2011). One explanation for this disconnect between theory and data is that environmental heterogeneity is rarely examined at sufficiently high resolution to evaluate its influence on within-community scale patterns of diversity (Adler et al. 2013; Kraft et al. 2014). More specifically, in observational studies, the local environment is

\*Correspondence author: E-mail: a.letten@unsw.edu.au

typically treated as spatially homogenous, with coexistence attributed to variation-independent processes such as resource partitioning, the temporal storage effect or neutral dynamics (Chesson 2000; Hubbell 2001). Matching this missing fine-scale environmental data with floristic observations will aid identification of the proximal causes of plant diversity maintenance at within-community scales.

Species coexistence arising through spatial heterogeneity is contingent on multiple criteria, but the most fundamental prerequisite of these is the presence of different species-specific responses to the environment (Chesson 2000; Silvertown 2004). More specifically, for a given spatially heterogeneous environmental variable to operate as a plant niche axis, and thus reduce niche overlap, species need to favour different regions of that variable. It follows that the spatial structure of the variable needs to exhibit sufficient temporal stability such that individuals may germinate, reach reproductive maturity and preferably build up a local population before the environment changes (Grubb 1977; Chesson 2000; Moko & Iwasa 2000; Amarasekare 2003). Of the vital plant resources, water in the form of soil moisture has repeatedly been shown to exhibit a high degree of temporal stability, whereby across fine scales, localities may often be consistently ranked on the basis of their soil water content (Vachaud et al. 1985; Cassel, Wendroth & Nielsen 2000; Pachepsky, Guber & Jacques 2005). As such, given the essential physiological role of water in plant function, and the vast array of plant adaptations to water stress and uptake, there is a strong *a priori* expectation that variability in soil moisture may provide one of the more important axes of niche differentiation at fine scales (Silvertown, Araya & Gowing 2015).

Whilst ecologists have long been aware of the tendency for plant species to segregate along strong soil moisture gradients (Pickett & Bazzaz 1978), evidence that subtle variation in soil moisture can also moderate co-occurrence patterns is only just emerging. Silvertown et al. (1999) first demonstrated fine-scale hydrological niche differentiation across topographically invariant landscapes in European wet meadows. More recently, Araya et al. (2011) underscored the potential generality of the phenomenon in *Restio*-dominated fynbos communities in the Cape Floristic region of South Africa. Given their combined taxonomic and geographic breadth, the evidence from these two studies is compelling, and yet there remains a strong mandate for wider investigation. In particular, we know very little about the comparative strength of soil moisture relative to other important plant resources (e.g. nitrogen, phosphorus and organic carbon) as drivers of fine-scale plant co-occurrence patterns.

Whilst manipulative experimental approaches arguably have the greatest potential to disentangle underlying mechanisms in annual plant communities (e.g. Sears & Chesson 2007; Godoy, Kraft & Levine 2014), they are less amenable to perennial systems where individuals can take several years to reach reproductive maturity. Fortunately, recent advances in statistical methods and computing power have greatly enhanced the capacity of researchers to draw robust inferences from the

kinds of purely observational data sets that are typical in studies of perennial communities. In particular, an emerging family of community-level model-based approaches has the potential to yield significant new insights (Ferrier & Guisan 2006; Warton et al. 2014). These include a subset of models which are specifically orientated towards disentangling patterns of co-occurrence and their abiotic and/or biotic drivers (Ovaskainen, Hottola & Siitonen 2010; Pollock et al. 2014; Harris 2015; Walker 2015). In contrast with conventional null-model based approaches, multivariate co-occurrence models (also known as joint species distribution models (JSDM), after Clark et al. 2013) provide a direct means of assessing shared environmental responses separately from other (abiotic and biotic) processes that may generate non-random patterns of co-occurrence. Furthermore, as they are built upon the standard regression framework of generalized linear models (McCullough & Nelder 1989), they facilitate more transparent and accurate representations of the statistical properties of the data (e.g. overdispersion of counts, Warton, Wright & Wang 2012), as well as provide a means of accounting for structural complexity and uncertainty in a hierarchical framework (Cressie et al. 2009).

In this study, we assessed the comparative importance of hydrological niche segregation for species co-occurrence, using a high-resolution study of soil hydrology and other edaphic variables, coupled with a long-term (24 years) data set of herbaceous plant plots in a coastal heathland community in south-east Australia. To this end, we modelled pairwise co-occurrence using an extension of the latent variable model (LVM, Skrandal & Rabe-Hesketh 2004) approach for model-based unconstrained ordination recently proposed by (Walker & Jackson 2011) and (Hui et al. 2014). LVMs offer an explicit, model-based approach to partitioning out the different drivers of species co-occurrence patterns. In particular, the regression component of an LVM models the species-specific environmental responses, based on which we can evaluate the prevalence of species-specific responses to soil hydrology, in comparison with a range of other important plant resources and non-resource factors. The latent variable component of the LVM is used to identify the residual patterns of species co-occurrence, which may be attributable to unmeasured factors and/or biotic interactions.

If heterogeneity in fine-scale hydrology is an important niche parameter, we predicted species would exhibit strongly differentiated responses to the local soil moisture gradient and that relative to other environmental variables, this would translate into large numbers of negative correlations between species. To evaluate the overall influence of the environment on patterns of species co-occurrence, we further examined the residual patterns of co-occurrence in an LVM including multiple environmental predictors in a single model. If the environment is the dominant force structuring patterns of negative species co-occurrence, we expected weak residual correlations induced by the latent variables. Conversely, if any unmeasured environmental factors and/or biotic processes are equally, or more, important than measured environmental

factors, we expected strong correlations based on the latent variables.

Given that observations of species distributions at fine spatial grains are known to be sensitive to stochastic processes (Chase 2014), we performed our primary analysis on all plots from all censuses combined, whilst accounting for spatial and temporal non-independence in a hierarchical framework. This aggregation of census data collected over a long period of time not only served to minimize the masking affect of stochastic phenomena on underlying species-environment associations, but also served to reduce measurement error that might arise from missed detections in any given single survey. Nevertheless, to aid interpretation and disentangle spatial and temporal processes, we additionally examined species-specific responses at each individual census, as well as at a finer spatial scale to the main analysis.

## Materials and methods

### STUDY AREA AND FLORISTIC SAMPLING

Sampling was conducted in an area (approximately 4 ha) of fire-prone coastal heathland in Royal National Park, New South Wales, Australia. The herbaceous ground layer is dominated by species within the Restionaceae, Cyperaceae and Poaceae families, whilst the overstorey consists mainly of sclerophyllous shrubs within the Proteaceae, Myrtaceae, Ericaceae and Fabaceae families (Keith et al. 2007). The soils derive from a sandstone substrate and tend to be low in nutrients, sandy and acidic. The topography is relatively flat, with minimum and maximum elevations of 67 and 71 m, respectively. A previous study found that terrain-based hydrological models provided poor estimates of true soil moisture at the site (Holman 1999), as is typical of low relief areas (Anderson & Kneale 1980).

Fifty-six 0.25 m<sup>2</sup> plots within the study area were sampled 10 times over 24 years (1990–1994, 1999, 2002, 2007, 2011 and 2014). At each census, the abundance (number of stems) of all herbaceous species in each plot was recorded. The plots are spatially clustered into four groups of 14, with a mean distance of 211 m between each cluster (min = 104 m; max = 323 m). Within each cluster, plots are spaced a mean distance of 18 m apart (min = 5 m, max = 45 m). A fire in October 1988 burnt the entire site prior to the commencement of the study, with subsequent fires in January 1994 (all plots burnt) and January 2001 (14 plots in one cluster burnt).

We restricted our analysis to the most dominant species, where dominance was defined as those species present in at least 20% of all 560 observations (56 plots × 10 censuses). Twelve species met this criterion. The rationale for only considering the dominant species was both ecological and methodological. Firstly, we asserted *a priori* that for any given environmental variable to be considered an important niche axis in the community, at least two of the dominant members should exhibit distinct responses to that variable. Secondly, the model-based framework we employed precluded the analysis of rare species for which too few data points were available.

### ENVIRONMENTAL SAMPLING

In March–May 2013, 56 100 cm access tubes for a PR2/6 soil moisture profile probe (Delta-T Devices: <http://www.delta-t.co.uk>) were installed within 10 cm of each plot. The probe is fitted with six sen-

sors allowing for instantaneous soil moisture measurements at six depths (10/20/30/40/60/100 cm). After a 2-month settling-in period, soil moisture sampling commenced in early August 2013 and was repeated at monthly intervals until July 2014. At each monthly interval, three readings (× six depths) were taken at each plot, with the probe rotated through 120° between each reading to account for any small-scale variability in moisture at each depth.

In addition to present-day measurement of soil moisture, a range of edaphic variables known to be important for plant growth and survival has been sampled in the immediate vicinity of each plot on seven occasions spread across the sampling period (1991, 1993, 1994, 1995, 1999, 2002 and 2014). At each plot, 5–10 soil cores (diameter: 27 mm; length = 100 mm) were sampled randomly within 10 cm of the plot boundary. In each given sampling year, the cores from each plot were combined and analysed for a range of variables, including organic carbon (all years), ammonium nitrogen (NH<sub>4</sub>) (all years excluding 1991, 1999 and 2014), nitrate nitrogen (NO<sub>3</sub>) (all years excluding 1991), total iron (all years excluding 1991, 1995 and 2002), phosphorus (all years excluding 1991, 1995 and 2002), conductivity (all years), pH (all years), cation exchange capacity (CEC) (all years) and various exchangeable cations (Na, K, Ca, Mg & Al) (all years).

### DATA PREPARATION

Prior to inclusion in models, soil variables were evaluated for several criteria, including sufficient spatial heterogeneity, relative temporal stability (a prerequisite for a spatial niche axis) and the absence of outliers. Both nitrate and exchangeable K were at such low levels that very little variability in their concentration could be detected and were consequently excluded from the analysis. Of the remaining variables, nitrogen (ammonium), conductivity, exchangeable Na and exchangeable Ca were found to exhibit substantial temporal variability with rapid temporal decay in their spatial correlation (mean Spearman correlation coefficient between sample years = 0.23, 0.17, 0.20 & 0.32, respectively) and so were also excluded. The mean value was then calculated for each of the remaining variables at each plot, following which 1–2 outliers were excluded for several variables to avoid them having a disproportional influence on parameter estimates. The final set of variables included phosphorus, organic carbon, total iron, pH, cation exchange capacity, exchangeable Mg and exchangeable Al, as well as two measures of soil moisture detailed further below.

In the absence of between-year samples of soil moisture at each plot, the temporal stability of soil moisture at each depth was evaluated on the basis of the variability observed over the 12-month sampling period. The vertical profile (100, 200, 300, 400, 600 and 1000 mm) for soil moisture exhibited a gradual decay in correlation such that values at 100 mm were near independent of those at 1000 mm ( $r = 0.09$ ), whilst those at intermediate depth fell along a spectrum between these two extremes. As a result, we subsequently restricted further evaluation of soil moisture to these two depths, one representing 'shallow' near-surface soil moisture (100 mm) and one representing 'deep' soil moisture (1000 mm). To evaluate the temporal stability of soil moisture, both in terms of relative spatial structure and absolute levels, we considered the mean Spearman rank correlation coefficient of all pairwise monthly comparisons (relative stability) and the overall standard deviation in mean monthly soil moisture (absolute stability). Deep soil moisture was highly stable in both relative and absolute terms, with mean pairwise correlation between months = 0.96 and standard deviation in average monthly soil moisture = 1.017. Shallow soil moisture was only slightly less stable with

mean pairwise correlation between months = 0.86 and standard deviation in average monthly soil moisture = 4.7. The high Spearman rank coefficients observed at both depths indicate high temporal stability in the relative spatial structure of soil moisture at the study site, whilst the increase in stability with depth is consistent with previous work (Cassel et al. 2000). As such there is strong evidence that the spatial structure of soil moisture observed over the 12-month monitoring period is typical of that which would have been observed throughout the floristic sampling period. Furthermore, variability in monthly rainfall over the 12-month monitoring period ( $SD = 57.3$ ) was near that observed for the entire period since floristic surveys began in 1990 ( $SD = 78.4$ ). The final per plot measurements of deep and shallow soil moisture were obtained by averaging across all 12 monthly samples. Notably, average shallow soil moisture was unrelated to elevation ( $P = 0.832$ ), whilst deep soil moisture was weakly and positively related ( $R^2 = 0.1031$   $P = 0.0179$ ), indicating a very slight increase in deep soil moisture with increasing elevation.

Finally, to obtain an estimate of light availability in the final year of sampling, bottom-up canopy photographs were taken over 2 days in March 2014. At each plot, two replicate photographs were taken using an SLR camera with a fish-eye lens positioned facing directly upwards on a level surface in an east–west orientation. The images were then digitally converted to black and white based on an automatic threshold using the GIMP graphics editor (<http://www.gimp.org>), and the ratio of black to white pixels (representing vegetation cover) was quantified. The two counts for each plot were then averaged.

## MODEL DESIGN

We used a novel approach to analysing co-occurrence patterns based on latent variable models (LVM). LVMs can be regarded as an extension of factor analysis (Knott & Bartholomew 1999) to non-normally distributed responses. For each of the nine environmental predictors (excl. light availability), we ran three different LVMs: one leveraging all plots and censuses combined (full model; Fig. 1a); one for each of the four clusters of plots (individual cluster model; Fig. 1b) to infer to what extent patterns observed for the full data set were driven by variation within or between clusters; and one for each of the ten individual censuses (individual census model; Fig. 1c) to evaluate the temporal stability of pairwise co-occurrence patterns. Additionally, we fitted an LVM for the full data set which included multiple environmental predictors (excl. light availability) in order to identify remaining patterns of co-occurrence after accounting for the environment. Due to collinearity ( $r > 0.7$ ) amongst several variables, three variables (pH, CEC and exchangeable Mg) were left out of the multipredictor model. We also fitted a single LVM for species co-occurrence as a function of light availability for 2014 only. For the individual cluster models, we restricted our analysis to those species that occurred in at least 20% of all plots in each cluster over the 10 censuses (as was done for the full model). For the individual census point models, the criterion was relaxed to 10% of all plots at each time step. A schematic summarizing the relationship between the different models and associated analyses (detailed further below) is provided in Fig. 1.

Similar to other co-occurrence models such as JSDMs (e.g. Ovaskainen, Hottola & Siitonen 2010; Clark et al. 2013; Pollock et al. 2014), the extent to which species exhibit distinct environmental responses was inferred through a fitted regression model, specifically using the framework of generalized linear models. However, instead of employing an unstructured covariance matrix to account for the missing environmental covariates and/or biotic interactions (as was

done in Ovaskainen, Hottola & Siitonen 2010; Pollock et al. 2014), LVMs utilize latent variables as a parsimonious means of modelling residual species correlation (see also Harris 2015). As such, we use the term ‘residual’ in reference to remaining patterns in the data after accounting for one or more predictors, rather than in terms of its definition in the context of residual analysis.

Our model can be regarded as an extension of the LVM proposed for model-based unconstrained ordination in Hui et al. (2014) and can be written in the following hierarchical form:

$$\text{Responses: } [y_{ij} | \mathbf{u}_i, \mathbf{x}_i] \sim \text{Neg-Bin}(y_{ij}; \mu_{ij}, \phi_j)$$

$$\log(\mu_{ij}) = \eta_{ij} = \mathbf{x}_i' \boldsymbol{\beta}_j + \mathbf{u}_i' \boldsymbol{\lambda}_j$$

$$\text{Latent Variables: } [\mathbf{u}_i] \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

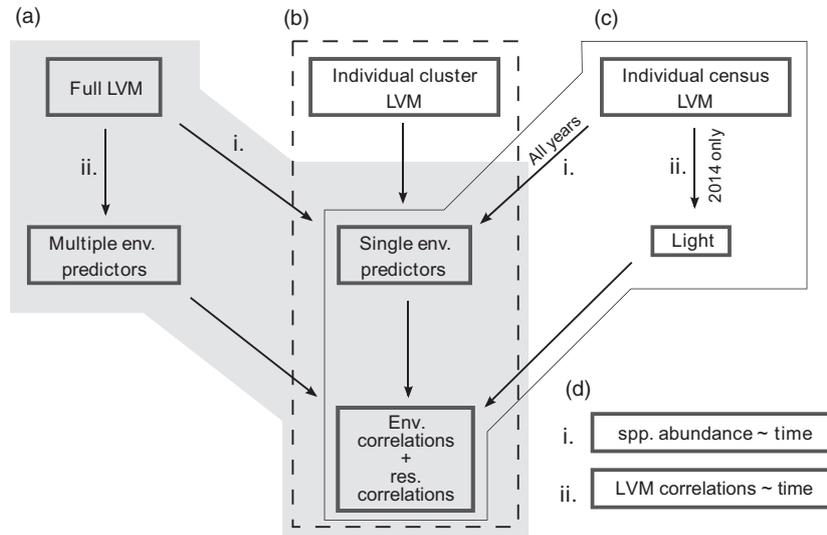
$$\text{Priors: } [\boldsymbol{\beta}_j] \sim \mathcal{N}(\mathbf{0}, c_0 \mathbf{I}), [\boldsymbol{\lambda}_j] \sim \mathcal{N}(\mathbf{0}, c_0 \mathbf{I}), [\phi_j] \sim \text{Unif}(0, c_1),$$

(1)

where ‘ $\sim$ ’ denotes ‘is distributed as’,  $\mathcal{N}(\cdot, \cdot)$  denotes a multivariate normal distribution with mean and covariance matrix given by the first and second arguments respectively,  $\text{Unif}(0, c_1)$  denotes a uniform distribution with minimum zero and maximum  $c_1$ , and  $\mathbf{I}$  denotes an identity matrix. To elaborate,  $y_{ij}$  denotes the observed count for species  $j$  at site  $i$  and is assumed to come from a negative binomial (Neg-Bin) distribution with mean  $\mu_{ij}$  and species-specific overdispersion parameter  $\phi_j$ . We used a negative binomial distribution as it exhibits a quadratic mean variance relationship,  $\text{Var}(y_{ij}) = \mu_{ij} + \phi_j \mu_{ij}^2$ , which allowed us to explicitly account for overdispersion present in the species counts.

Using a log-link function, the mean  $\mu_{ij}$  was regressed against two sets of variables. Firstly, a vector of explanatory variables for each site,  $\mathbf{x}_i$ , which included an intercept term, environmental covariates (e.g. soil moisture), a fixed effect for the four clusters of plots to account for spatial non-independence within clusters (full and individual year models) and a random effect (intercept) for plot to account for temporal non-independence (full and individual cluster models). In order to account for unimodal species responses as predicted by niche theory (Austin 2002), quadratic terms were included for the environmental covariates in each model. The second set of variables comprised a vector of two latent variables for each site,  $\mathbf{u}_i = (u_{i1}, u_{i2})$ , which were assumed to be drawn from independent, standard normal distributions. The species-specific regression coefficients  $\boldsymbol{\beta}_j$  and  $\boldsymbol{\lambda}_j$  describe how the mean changes as a function of the explanatory and latent variables, respectively.

The key element in the LVM in eqn (1) is the latent variables  $\mathbf{u}_i$ , which can account for any residual correlation between species not attributable to spatial heterogeneity in the measured environmental covariates  $\mathbf{x}_i$ . This correlation may be driven by biotic interactions such as competition (negative) or facilitation (positive) or alternatively to missing predictors, where  $\boldsymbol{\lambda}_j$  is the coefficient corresponding to these missing predictors. In addition, in our full model which combines multiple censuses, residual correlation may arise due to negatively or positively correlated fluctuations in species abundance through time. If unaccounted for, missing covariates, ecological interactions or temporal correlation will induce residual correlation between species which can potentially lead to erroneous inference. Latent variables offer an attractive approach to dealing with this issue. In particular, they require significantly less parameters to model species residual correlation compared to the unstructured (full rank) correlation matrices used by Ovaskainen, Hottola & Siitonen (2010) and Pollock et al. (2014). Finally, note that, more than two latent variables could have been used, although our preliminary testing with this data set suggested that two was sufficient to characterize the main patterns of species residual correlations.



**Fig. 1.** Schematic summary of core analysis. Environmental and residual correlations calculated for: (a) LVMs fitted to the full data set for i) each of the nine environmental predictors independently and ii) for multiple predictors in a single model; (b) LVMs fitted to each individual cluster for each predictor independently; and (c) i) LVMs fitted to each census for each predictor independently (all census years) and ii) additionally for light in 2014. Supplementary GAMMs were fitted for (d) i) species abundance as a function of time for comparison with the residual correlations from the multiple predictor model and ii) for environmental and residual correlations from the individual census LVMs as a function of time.

We used Bayesian Markov Chain Monte Carlo (MCMC) methods to estimate the LVMs, with sampling performed through JAGS v3.4.0 (Plummer 2003) using the package R2jags v0.03-08 (Su & Yajima 2012) in R v3.1.1. We assigned uninformative priors for all parameters,  $\beta_j$ ,  $\lambda_j$  and  $\phi_j$ , by setting  $c_0 = c_1 = 100$  in eqn (1). For the full and individual cluster models, we also assigned an uninformative uniform prior for the variances of the random effect for plot. For each LVM fitted, we ran three chains with a burnin period of 10 000 iterations followed by 100 000 iterations with a thinning lag of 50 for each chain. This produced a final combined sample of 6000 MCMC samples for each LVM. We assessed parameters to have converged when traceplots were well mixed and the Gelman–Rubin statistic was below 1.1.

R code demonstrating the fitting and analysis of latent variable models of co-occurrence is provided in Appendix S1 of the Supporting Information.

#### MODEL INFERENCE AND SUPPLEMENTARY ANALYSIS

After fitting the LVMs, in order to visualize patterns of co-occurrence arising from the different environmental factors, we calculated two types of correlation matrices. The first was constructed by calculating, for any two species, the correlation between their fitted values  $x_i^j \beta_j$  (across all plots). This is the same as eqn (4) in Pollock et al. (2014), with the resulting matrix representing the correlation between species that can be attributed to a shared/diverging environmental response. The second type of correlation matrix was calculated using the latent variable coefficients,  $\lambda_j$ , also known as factor loadings. Specifically, if we let  $\Lambda$  be the two-column matrix formed by stacking the factor loadings on top of each other, then a covariance matrix is obtained as  $\Lambda \Lambda'$  from which the residual correlation matrix can be calculated. This second residual correlation matrix represents the correlation between species that may be attributable to biotic interactions or missing environmental covariates. Since Bayesian MCMC estimation was used, the correlation between fitted responses was calculated for each MCMC sample, which made it possible to obtain a posterior distribution for each cell of the environmental and residual correlation matrix.

As such, correlation ‘significance’ was evaluated on the basis of the 95% credible intervals for the posterior mean excluding zero.

Having extracted environmental and residual correlation matrices for each environmental factor in each of the three model types (full, individual cluster and individual census point), we assessed the importance of each factor as a niche axis on the basis of two criteria: (i) the number of significant negative pairwise environmental responses, where significance is defined as 95% credible intervals that do not cross zero; and (ii) the minimum number of species that need to be excluded to remove all negative pairwise responses, referred to in network theory as the minimum vertex cover. In this application, the minimum vertex cover quantifies the extent to which a given number  $n$  of negative pairwise environmental responses is driven by one species exhibiting a different response to all other species, or at the other extreme consists of completely unique species pairs. In the former case, only one species need to be excluded to remove all negative correlations, whereas in the latter case,  $n$  species must be removed. As such, with 12 species, if all 66 unique pairs exhibit significant negative correlations (an implausible scenario necessitating extremely narrow niche breadths), 11 would need to be excluded. In addition to quantifying negative pairwise environmental responses, we also counted the number of positive (shared) environmental responses to evaluate the extent to which a given environmental factor influences the fine-scale distribution of species even in the absence of any clear niche differentiation (e.g. if all species exhibit a similar response to the gradient).

After considering species-specific environmental responses, we then examined the residual correlation matrix for each model, with particular emphasis on the full model including all of the environmental predictors, as this represents patterns of co-occurrence unexplained by all our measured predictors combined. To aid interpretation of the residual correlation matrix from the full multipredictor model, we also fitted a generalized additive mixed models (GAMMs) for each species as a function of time and compared the correlation in the fitted values from the GAMMs with the residual correlation matrix based on the full LVM model including multiple environmental predictors. The temporal GAMMs for each species were fitted using the MGCV package (Wood 2012), with a thin-plate spline smoother for time, a

fixed effect for plot cluster, a random effect for plot, an AR(1) correlation structure nested within plot and assuming a negative binomial error distribution with log link. If patterns in the residual correlation matrix of the full model are largely a product of temporal correlation (negative or positive) in species responses, we expected a strong correlation with the fitted value correlation matrix derived from the individual GAMMs. Correlation between the two matrices was assessed via a Mantel test. Note that, the GAMM for one species, *Empodisma minus*, failed to converge (most likely due to overdispersion) and as a result had to be excluded from the temporal analyses.

To investigate temporal trends in environmental and residual correlations, we also fitted GAMMs for mean pairwise environmental and residual correlation as a function of time for all variables combined. Each model included a thin-plate spline smoother for time, a random effect for environmental variable and an AR(1) correlation structure nested within environmental variable. Owing to the mean correlations values being consistently positive and right skewed, we used a gamma error distribution with a log-link function. To correct over-smoothing observed in the residual correlation trends model, the maximum number of degrees of freedom was reduced from 9 (package default) to 4.

Finally, we performed *post hoc* exploratory analysis of the differences in species response to soil moisture with their pairwise differences in two readily available aboveground traits, plant height and seed mass. In each case, we used a Mantel test to compare pairwise difference in soil moisture ( $-1 \times$  environmental correlation) with pairwise trait differences and visualized the relationship with scatter plots.

## Results

Of the nine environmental variables considered independently in each of the single-predictor models for the full data set, all but pH produced significant negative correlations due to diverging species-specific responses (Fig. 2). Deep soil moisture was associated with both the largest number (18) of negative correlations and the largest number (4) of node removals required to eliminate all negative correlations. As is to be expected given such strong correlations, the 12 species displayed a range of differentiated responses to the deep soil moisture gradient, with some increasing monotonically (or near monotonically) towards the extremes of the gradient (e.g. *Xanthorrhoea resinosa*, *Haemodorum corymbosum* and *Entolasia stricta*), whilst others exhibited unimodal peaks closer to the middle of the gradient (e.g. *Ptilothrix deusta*, *Empodisma minus* and *Xyris gracilis*) (Fig. 3). Exchangeable aluminium was associated with the second largest number (17) of negative correlations, followed by organic carbon (11), shallow soil moisture and total iron (10), exchangeable magnesium (6), cation exchange capacity (3) and phosphorus (1) (Fig. 2). Whilst pH was not associated with any negative pairwise correlations, 27 species pairs exhibited significant positive correlations to pH. Deep soil moisture was associated with the second highest number (25) of significant positive correlations, followed by exchangeable aluminium and exchangeable magnesium (21), organic carbon (18), shallow soil moisture (17), total iron (16), cation exchange capacity (12) and phosphorus (8) (see Fig. S1 in Supporting Information).

Residual correlations based on single-predictor LVMs fitted to the full data set were strongly associated with each other and

with residual correlations from the multipredictor model ( $r = 0.68\text{--}0.97$ ). As such, the consistency of the observed residual correlations suggests a potentially important variable was not accounted for in any of the LVMs. Nevertheless, most of the significant correlations in the residual correlation matrix of the multipredictor model were positive (23) rather than negative (2), thus negating the probability that an important fine-scale niche axis was omitted from our analysis (Fig. 4). Furthermore, a substantial portion of the residual correlation appears to be attributable to shared and differentiated temporal responses between species, with the residual correlation matrix from the multipredictor model being moderately associated with the temporal response correlation matrix derived from the temporal models (Mantel test  $r = 0.43$ ) (see Fig. S2 for smoothed species responses through time). Note that, as a likely consequence of large overdispersion, the temporal model for one species, *Empodisma minus*, failed to converge.

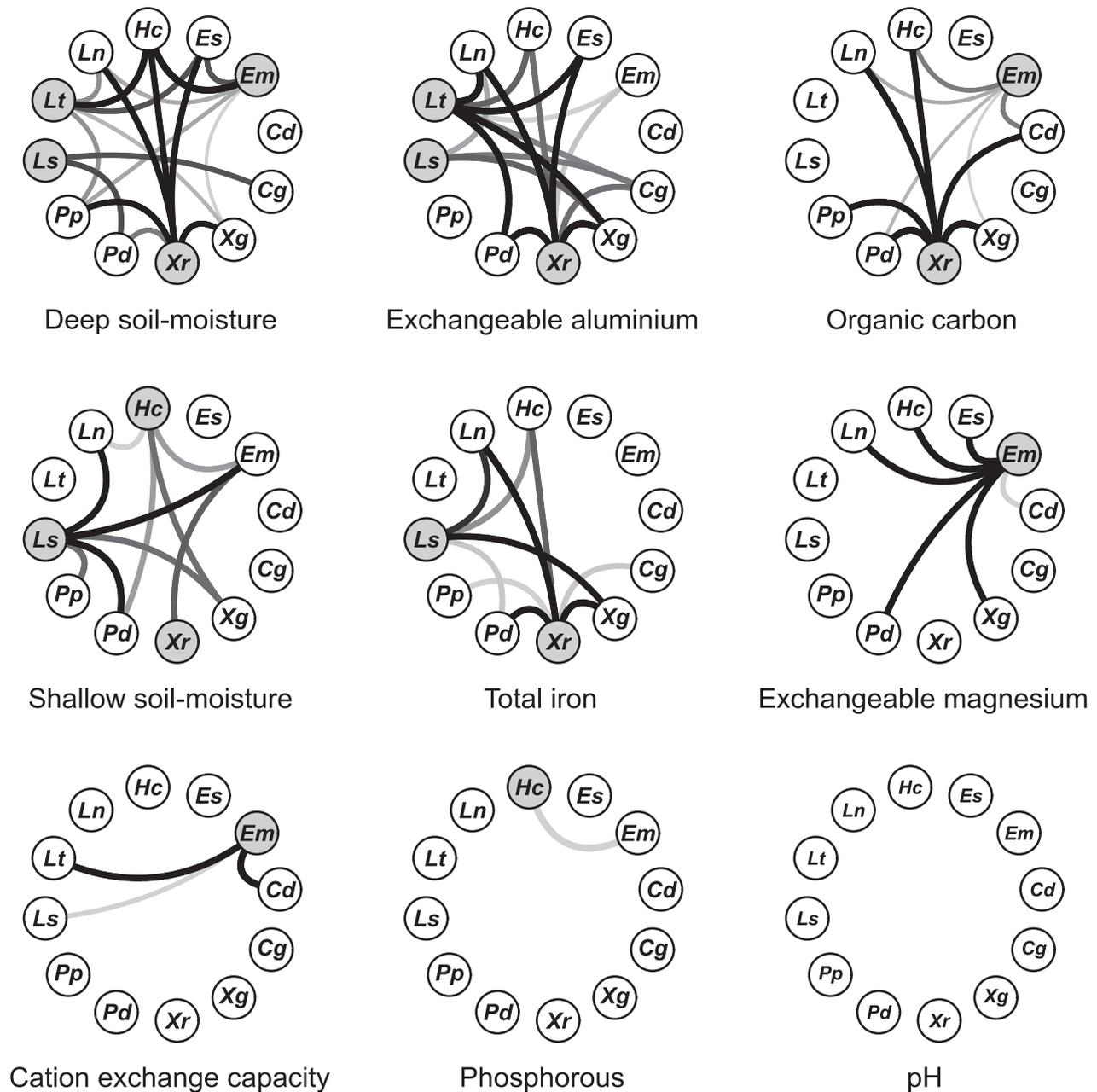
Relative to the models run on the full data set, the number of significant negative correlations due to diverging environmental responses between species was considerably smaller for models run on each individual cluster. As such, most of the observed patterns of co-occurrence in the full models appear to be driven by environmental variation over spatial scales of 10s of metres, rather than metres (i.e. between clusters rather than within clusters). The one exception to this pattern was for phosphorus, which exhibited a relatively high number of significant negative co-occurrence associations in three of the four clusters. Notably, a substantial number of species did exhibit negative pairwise associations in the residual correlation matrices for at least two of the clusters, suggesting that biotic interactions may still influence co-occurrence at within-cluster scales (see Table S1).

The number of significant negative pairwise associations in species environmental response was also comparatively small for the individual census models. It is notable, however, that even if we ignore 'significance' criteria, the sign of each pairwise species association is largely consistent with those observed for the single-predictor models for the full data set. In addition, the models for each individual census provide insight into the relative dominance of positive and negative interactions through time. Whilst there were no obvious trends in mean environmental correlation through time, residual correlations underwent a conspicuous upwards positive trend in the second half of the sampling period (Fig. 5). At the same time, pairwise species associations derived from the single LVM fitted for light availability in 2014 were almost universally positive, with all but one species (*Xanthorrhoea resinosa*) being more abundant in less shaded locations.

Differences in species pairwise responses to deep soil moisture were moderately positively related to pairwise difference in plant height ( $r = 0.6259$ ,  $P = 0.002$ ) but unrelated to differences in seed mass ( $r = -0.1176$ ,  $P = 0.809$ ) (Fig. 6).

## Discussion

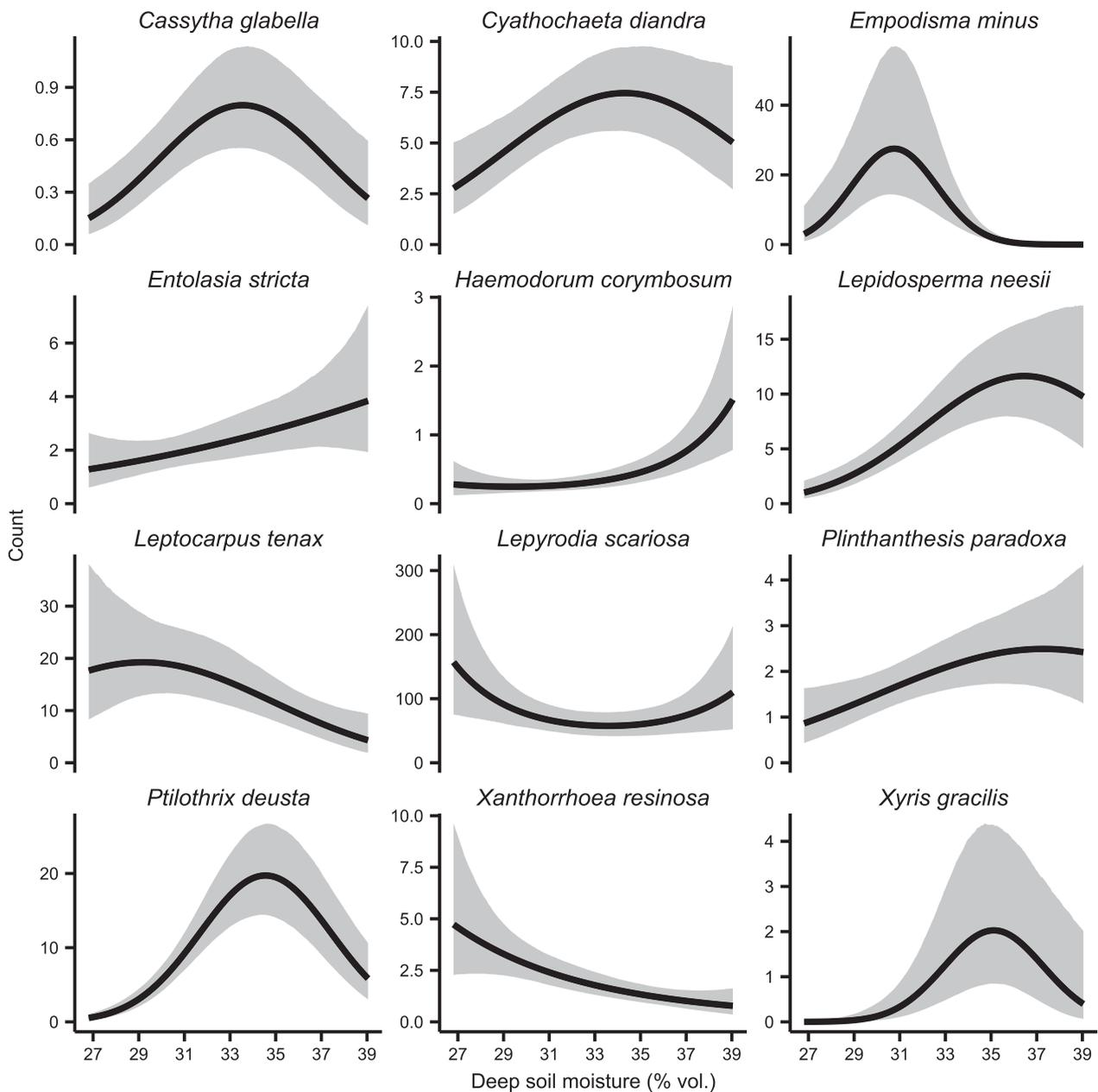
The notion that fine-scale spatial environmental variability may foster plant species coexistence is founded upon a rich



**Fig. 2.** Negative pairwise species correlations to the environment derived from single-predictor LVMs for the full data set. Connecting lines between species nodes denote negative mean posterior correlations with credible intervals excluding zero. Line colour and thickness indicate the strength of the negative correlation where darker and thicker lines are closer to  $-1$ . Nodes shaded grey indicate the minimum vertex cover, that is the smallest combination of species that need to be removed to break all negative associations (note that, whilst the minimum number has a single solution, the species composition making up the minimum vertex cover set can vary). Species node labels combine the first letters of the genus and specific epithet given in full in Fig. 3.

body of theory (Chesson 2000; Amarasekare 2003; Adler et al. 2013), and yet empirical efforts to establish which, if any, environmental factors are associated with strong patterns of fine-scale niche differentiation are surprisingly sparse. We found strong empirical support for the hypothesis that differentiation along hydrological gradients represents one of the most important fine-scale plant niche axes at within-community scales (*sensu* Silvertown, Araya & Gowing 2015). On the basis of multiple criteria, deep soil moisture emerged as the single best predictor of negative

co-occurrence patterns (Fig. 2), with the dominant community members exhibiting strongly differentiated responses across a comparatively narrow moisture gradient (Fig. 3). Together with several earlier studies (e.g. Silvertown et al. 1999; Araya et al. 2011), these results are amongst the first to demonstrate hydrological niche differentiation in the absence of significant topographical complexity. Furthermore, they highlight the robustness of the hypothesis when confronted with a comprehensive set of alternative prospective niche axes.

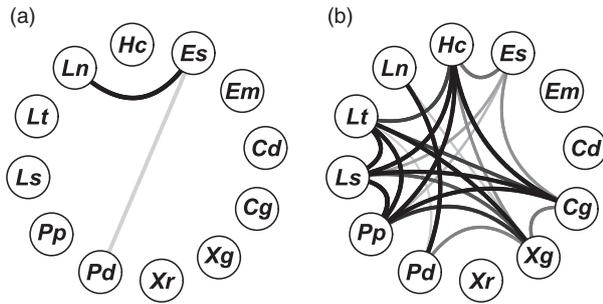


**Fig. 3.** Species-specific responses to the deep soil moisture gradient. Fitted values derived from regression parameter estimates from the deep soil moisture LVM (full data set). Shaded region denotes 95% credible intervals.

The adoption of a latent variable modelling framework enabled us not only to evaluate the comparative importance of each environmental factor, but also to determine the likelihood that an important factor was unaccounted for. Given that the residual correlation matrix from the multipredictor model was dominated by positive pairwise associations, we can be confident that we accounted for the most important factors driving niche segregation amongst the dominant community members (Fig. 4). As such, our approach allows robust conclusions about the relative importance of soil hydrology not only with respect to measured variables but also unmeasured variables. Notably, a considerable portion of the residual correlation from the multipredictor model appeared to be attributable to both positive, and to a lesser extent negative, temporal

correlations in species peak abundance. These in turn appear to largely reflect differential interspecific temporal responses to fire events in 1988, 1994 and 2001 (Fig. S2). As such, with time-series data, co-occurrence models may also be used to draw inferences on successional processes and temporal niche partitioning.

Whilst the results provide strong evidence for fine-scale hydrological niche differentiation, the extent to which hydrological niches provide opportunities for local-scale species coexistence is to a large extent contingent on the scale at which the community is defined (Amarasekare 2003; Adler et al. 2013). For instance, in the current study, negative co-occurrence patterns associated with the deep soil moisture gradient were only detectable when comparing plots across



**Fig. 4.** Negative (a) and positive (b) pairwise residual correlations derived from the multipredictor LVM for the full data set. Connecting lines between species nodes denote posterior correlations with credible intervals excluding zero. Line colour and thickness indicates the strength of the correlation where darker and thicker lines are closer to  $|1|$ . Species node labels combine the first letters of the genus and specific epithet.

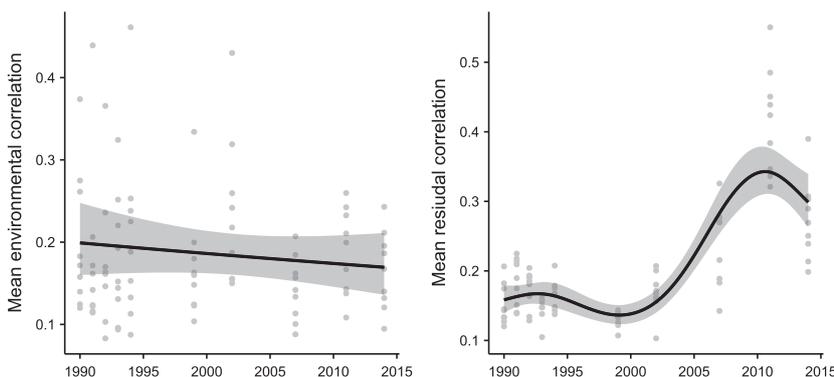
clusters separated by 100–320 m. Similarly, Silvertown et al. (1999) made what is probably the strongest case for hydrological niches to date on the basis of plots dispersed across 10's of hectares, whilst even the ca.  $50 \times 50$  m plots of Araya et al. (2011) presumably allowed for individuals to be spaced as much as 70 m apart. In contrast, Fridley et al. (2011) reported species-specific responses to microsite variation in soil depth in  $3 \times 3$  m plots in a limestone grassland, but this appeared to be unrelated to soil water potential. As such, the evidence for hydrological niches at the local scale within which neighbouring plants directly compete for resources arguably remains relatively weak. This conclusion would be in line with the dominant paradigm in community ecology that negative patterns of co-occurrence at the local scale are predominately indicative of competitive exclusion (Gotelli & McCabe 2002, but see Fridley et al. 2011). Indeed, only phosphorus exhibited more than six negative pairwise associations (compared with just one for the full data set) in any one plot cluster (Table S1). In contrast, the residual correlation matrices for two plot clusters exhibited a relatively large number of negative associations, which do not appear to be driven by any of the measured environmental factors. Nevertheless, even in the absence of localized spatial partitioning of resources, hydrological niches may still promote local coexistence via source-sink dynamics and spatial storage effects (Shmida & Ellner 1984; Amarasekare 2003; Snyder & Chesson 2004). This will arise when species are able to disperse

from favourable to less favourable locations at a sufficiently high rate to offset locally negative population growth rates and thus prevent local extinction (Snyder & Chesson 2004). Since monitoring began in the current study, all but one species has been recorded at all four clusters, suggesting that most of the dominant community members do indeed regularly disperse to and establish in locations where they are less competitive. It is also notable, that within short distances ( $< 100$  m) of our study plots, the topography is more varied, and as such may provide nearby opportunities for even finer scale hydrological niche differentiation.

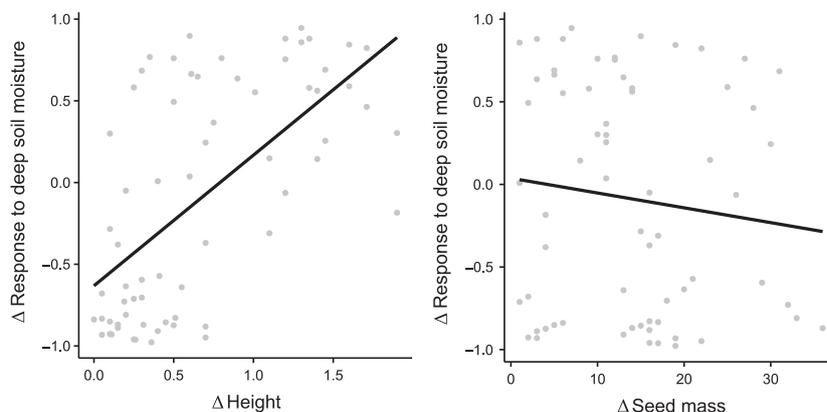
In contrast to the individual cluster models, the lack of strong negative environmental correlations in the individual census models appears to be an artefact of low power, particularly given that the raw correlation values corresponded closely with those from the full model. Nevertheless, it is interesting to note that whereas mean environmental correlations were relatively static through time, residuals correlations exhibited a strong upwards trend towards more positive values. This trend most likely reflects a shared environmental response to increased overstorey shading, as indicated by the observed positive response of the majority of species to light availability in the most recent 2014 census. In a previous study from the same site, Letten, Keith & Tozer (2014) described an increase in functional and phylogenetic similarity of community members through temporal succession consistent with this observation.

Our focus was on negative environmental correlations, but positive correlations also provide insight into other aspects of community assembly. For instance, positive correlations in species responses were highest for soil pH, suggesting pH may be a useful predictor of local richness patterns, whilst also acting as a niche axis over larger spatial scales. It is also notable that although this system is typically associated with low-pH adapted species, many of the dominant herbaceous community members exhibited a positive response to the pH gradient (data not shown). As such, it would be interesting to explore the extent to which rarer species exploit this apparently 'unoccupied' niche space within the landscape.

Although a comprehensive examination of trait–environment relationships was beyond the scope of this study, comparisons of species pairwise differences in seed mass and plant height relative to hydrological niches differences still



**Fig. 5.** Trends in mean environmental (left) and residual (right) correlation through time. Trend lines obtained with thin-plate splines in a GAMM framework (see main text for model fitting). Shaded region represents 95% confidence intervals.



**Fig. 6.** Scatter plots of pairwise differences in species responses to deep soil moisture ( $-1 \times$  environmental correlation) in relation to pairwise difference in plant height (left) and seed mass (right). Lines of best fit obtained from linear models.

proved informative. Whereas differences in seed mass, which may be associated with survival from drought stress (Pérez-Harguindeguy et al. 2013), were unrelated to hydrological niches, differences in plant height were strongly positively related with differences in species response to soil moisture (Fig. 6; cf. Fig 2c in Adler et al. 2013). This appeared to be driven predominately by taller species (e.g. *Leptocarpus tenax* and *Xanthorrhoea resinosa*) favouring drier areas of the deep soil moisture gradient, whilst shorter species (e.g. *Plinthanthesis paradoxa* and *Xyris gracilis*) were more abundant in wetter areas. Given that soil moisture measurements at all depths were typically around, or above, the field capacity of sandy-loam soil ( $\sim 23\%$  vol., Atwell, Kriedemann & Turnbull 1999), this possibly reflects the greater tolerance of shorter species to low oxygen in periodically saturated sites. In addition it suggests fine-scale variability in soil moisture may have important implications for the structural complexity of standing vegetation.

Beyond differences in plant height and seed mass, it is notable that the dominant herbaceous members of the community actually exhibit remarkably similar phenotypes, with all but one of the 12 study species being a geophytic or hemicryptophytic monocot and nine of the 12 belonging to just three families within the Poales (*Restionaceae*, *Poaceae* and *Cyperaceae*). This apparent constraint on functional strategies is likely attributable to the fire-prone and nutrient poor nature of the system, which makes it all the more remarkable that these same species exhibit such differentiated responses to the soil moisture gradient. One likely explanation is that these species exhibit a range of different below-ground adaptations to water acquisition and flooding/drought tolerance (Silvertown, Araya & Gowing 2015).

An open question arising from our results is why species co-occurrence patterns should exhibit a stronger relationship with deep, rather than shallow, soil moisture. This is particularly true given that most of the study species are likely to have the bulk of their rooting volume above 1000 mm where the deep soil moisture measurements were made. The most parsimonious explanation is that shallow soil moisture is unevenly modified by plant water usage and episodic rainfall recharge and thus provides a poorer approximation of baseline soil moisture across the study area. More specifically, shallow soil moisture is likely to experience greater seasonal and in-

terannual variability contingent on the density of standing vegetation. Observed trends in monthly soil moisture support this notion, with shallow soil moisture exhibiting a relatively strong seasonal trend over the 12 months of sampling, which appeared to be mostly attributable to increased plant water usage over spring/summer rather than decreased rainfall. In contrast, deep soil moisture was comparatively stable. This may partly explain the apparent lack of a strong relationship between species-specific responses and near-surface soil moisture observed by Fridley et al. (2011). Ultimately, the depth of soil moisture that best predicts hydrological niches is likely to depend on numerous system-specific factors, but our findings highlight the potentially confounding role of plant water usage on inferences drawn from near-surface measurements. Furthermore, it is notable that elevation was a poor predictor of soil moisture at any depth, confirming that the use of terrain-based hydrological models in plant niche studies may lead to misleading conclusions.

One challenge to interpreting plant compositional responses to spatial variability in soil moisture is the dual role of water as both a depletable resource factor (cf. nutrients or light) subject to density-dependent feedbacks and as a non-resource factor (cf. temperature or soil type) that mediates density-independent responses (e.g. via reduced oxygen diffusion). It follows that depending on supply rates, spatial heterogeneity in soil moisture may produce complex interactions between niche and fitness differences amongst interacting species (*sensu* Chesson 2000). It might therefore be expected that the breadth of plant responses to soil moisture will allow for higher dimensional trade-offs relative to more narrowly defined environmental factors (Tilman 1982; Chase & Leibold 2003). In the current study, soil moisture appeared to be consistently around or above the likely field capacity of the soil. This suggests that in this system soil moisture may exert a stronger influence on compositional patterns by acting as a stressor rather than a spatially heterogeneous limiting resource. In the absence of data on spatial variation in air-filled porosity, such inferences are unfortunately, somewhat speculative, but this represents an important consideration for future studies.

It is also important to recognize that demonstrating species-specific environmental responses along a given gradient is not on its own sufficient evidence that it facilitates species coexis-

tence. To this end, rigorous tests of species coexistence necessitate experimental manipulations of intra- and interspecific competition (e.g. Levine & Rees 2002; Sears & Chesson 2007; Godoy, Kraft & Levine 2014), or alternatively the mathematical parametrization of models from long-term demographic studies (e.g. Adler et al. 2006; Angert et al. 2009). Unfortunately these approaches are not always practical, particularly in perennial systems such as heathlands where individuals can take several years to reach reproductive maturity. Nevertheless, as advocated by Silvertown, Araya & Gowing (2015), providing evidence of fine-scale hydrological niche differentiation is an important first step that will hopefully provide stimulus for further research.

Here, we have provided robust evidence for species-specific responses to fine-scale variation in soil hydrology, whilst also illustrating the valuable inferential gains made possible with model-based approaches to co-occurrence analysis. Our results are consistent with emerging empirical evidence, but go a step further than previous studies by demonstrating the apparent strength of hydrological niche differentiation when compared to a range of other potential niche axes. Future research efforts should focus on identifying the different physiological processes and traits associated with hydrological niches (Silvertown, Araya & Gowing 2015) and on exploring scale dependencies to better understand the underlying coexistence mechanisms at play (Adler et al. 2013). In the light of perturbations to hydrological regimes under climate change, a better understanding of hydrological niches could prove critical to conservation efforts.

## Acknowledgements

For their assistance with fieldwork, we are grateful to Chantel Benbow, Sam Dawson, Chris Gordon, Anna Feit, Ben Feit, Sylvia Hay, Mitch Lyons and Max Mallen-Cooper. The authors thank Jason Fridley and two anonymous reviewers for comments on an earlier draft of the manuscript and Martin Westgate whose 'circplot' R package was used to produce the network plots in Figs. 2 and 4.

## Data accessibility

Data to be made available in the Long Term Ecological Research Network Data Portal (<http://www.ltern.org.au/>) mid-2015.

## References

- Adler, P.B., Fajardo, A., Kleinhesselink, A.R. & Kraft, N.J.B. (2013) Trait-based tests of coexistence mechanisms. *Ecology Letters*, **16**, 1294–1306.
- Adler, P.B., HilleRisLambers, J., Kyriakidis, P.C., Guan, Q. & Levine, J.M. (2006) Climate variability has a stabilizing effect on the coexistence of prairie grasses. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 12793–12798.
- Amarasekare, P. (2003) Competitive coexistence in spatially structured environments: a synthesis. *Ecology Letters*, **6**, 1109–1122.
- Anderson, M. & Kneale, P. (1980) Topography and hillslope soil water relationships in a catchment of low relief. *Journal of Hydrology*, **47**, 115–128.
- Angert, A.L., Huxman, T.E., Chesson, P. & Venable, D.L. (2009) Functional trade-offs determine species coexistence via the storage effect. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 11641–11645.
- Araya, Y.N., Silvertown, J., Gowing, D.J., McConway, K.J., Linder, H.P. & Midgley, G. (2011) A fundamental, eco-hydrological basis for niche segregation in plant communities. *The New Phytologist*, **189**, 253–258.
- Atwell, B.J., Kriedemann, P.E. & Turnbull, C.G. (1999) *Plants in Action: Adaptation in Nature, Performance in Cultivation*. Macmillan Education AU, Melbourne.
- Austin, M. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101–118.
- Cassel, D., Wendroth, O. & Nielsen, D. (2000) Assessing spatial variability in an agricultural experiment station field: opportunities arising from spatial dependence. *Agronomy Journal*, **92**, 706–714.
- Chase, J.M. (2014) Spatial scale resolves the niche versus neutral theory debate. *Journal of Vegetation Science*, **25**, 319–322.
- Chase, J. & Leibold, M. (2003) *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press, Chicago, IL.
- Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, **31**, 343–366.
- Clark, J.S., Gelfand, A.E., Woodall, C.W. & Zhu, K. (2013) More than the sum of the parts: forest climate response from joint species distribution models. *Ecological Applications*, **24**, 990–999.
- Cressie, N., Calder, C.A., Clark, J.S., Hoef, J.M.V. & Wikle, C.K. (2009) Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, **19**, 553–570.
- Ferrier, S. & Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, **43**, 393–404.
- Fridley, J.D., Grime, J.P., Askew, A.P., Moser, B. & Stevens, C.J. (2011) Soil heterogeneity buffers community response to climate change in species-rich grassland. *Global Change Biology*, **17**, 2002–2011.
- Godoy, O., Kraft, N.J.B. & Levine, J.M. (2014) Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology Letters*, **17**, 836–844.
- Gotelli, N.J. & McCabe, D.J. (2002) Species co-occurrence: a meta-analysis of J. M. Diamond's assembly rules model. *Ecology*, **83**, 2091–2096.
- Grubb, P.J. (1977) The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews*, **52**, 107–145.
- Harris, D.J. (2015) Generating realistic assemblages with a Joint Species Distribution Model. *Methods in Ecology and Evolution*, **6**, 465–473.
- Holman, L. (1999) *Spatial analysis of soil moisture, nutrient status and vegetation within a wet heathland*. Honours, University of Sydney.
- Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*, Vol. 17 of *Monographs in Population Biology*. Princeton University Press, Princeton, NJ.
- Hui, F.K., Taskinen, S., Pledger, S., Foster, S.D. & Warton, D.I. (2014) Model-based approaches to unconstrained ordination. *Methods in Ecology and Evolution*, **6**, 399–411.
- Keith, D.A., Holman, L., Rodoreda, S., Lemmon, J. & Bedward, M. (2007) Plant functional types can predict decade-scale changes in fire-prone vegetation. *African Journal of Ecology*, **95**, 1324–1337.
- Knott, M. & Bartholomew, D.J. (1999) *Latent Variable Models and Factor Analysis*. 7. Edward Arnold, London.
- Kraft, N.J.B., Adler, P.B., Godoy, O., James, E., Fuller, S. & Levine, J.M. (2014) Community assembly, coexistence, and the environmental filtering metaphor. *Functional Ecology*, **29**, 592–599.
- Letten, A.D., Keith, D.A. & Tozer, M.G. (2014) Phylogenetic and functional dissimilarity does not increase during temporal heathland succession. *Proceedings of the Royal Society. B, Biological Sciences*, **281**, 20142102.
- Levine, J.M. & Rees, M. (2002) Coexistence and relative abundance in annual plant assemblages: the roles of competition and colonization. *The American Naturalist*, **160**, 452–467.
- McCullough, P. & Nelder, J.A. (1989) *Generalized Linear Models*. Chapman & Hall/CRC, Boca Raton, FL.
- Muko, S. & Iwasa, Y. (2000) Species coexistence by permanent spatial heterogeneity in a lottery model. *Theoretical Population Biology*, **57**, 273–284.
- Ovaskainen, O., Hottola, J. & Siitonen, J. (2010) Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology*, **91**, 2514–2521.
- Pachepsky, Y.A., Guber, A.K. & Jacques, D. (2005) Temporal persistence in vertical distributions of soil moisture contents. *Soil Science Society of America Journal*, **69**, 347–352.
- Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jauregui-berry, P. et al. (2013) New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany*, **61**, 167.
- Pickett, S.T.A. & Bazzaz, F.A. (1978) Organization of an assemblage of early successional species on a soil moisture gradient. *Ecology*, **59**, 1248.
- Plummer, M. (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March, pp. 20–22.

- Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A. & McCarthy, M.A. (2014) Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, **5**, 397–406.
- Sears, A.L.W. & Chesson, P. (2007) New methods for quantifying the spatial storage effect: an illustration with desert annuals. *Ecology*, **88**, 2240–2247.
- Shmida, A. & Ellner, S. (1984) Coexistence of plant species with similar niches. *Vegetatio*, **58**, 29–55.
- Silvertown, J. (2004) Plant coexistence and the niche. *Trends in Ecology & Evolution*, **19**, 605–611.
- Silvertown, J., Araya, Y. & Gowing, D. (2015) Hydrological niches in terrestrial plant communities: a review. *African Journal of Ecology*, **103**, 93–108.
- Silvertown, J., Dodd, M.E., Gowing, D.J.G. & Mountford, J.O. (1999) Hydrologically defined niches reveal a basis for species richness in plant communities. *Nature*, **400**, 61–63.
- Skrondal, A. & Rabe-Hesketh, S. (2004) *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. CRC Press, Boca Raton, FL.
- Snyder, R.E. & Chesson, P. (2004) How the spatial scales of dispersal, competition, and environmental heterogeneity interact to affect coexistence. *The American Naturalist* **164**, 633–650.
- Stein, A., Gerstner, K. & Kreft, H. (2014) Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, **17**, 866–80.
- Su, Y.S. & Yajima, M. (2012) *R2jags: A Package for Running jags from R*. R package version 0.03-08, available at <http://CRAN.R-project.org/package=R2jags>.
- Tilman, D. (1982) *Resource Competition and Community Structure (Mpb-17)*. Princeton University Press, Princeton, NJ.
- Vachaud, G., Passerat De Silans, A., Balabanis, P. & Vauclin, M. (1985) Temporal stability of spatially measured soil water probability density function. *Soil Science Society of America Journal*, **49**, 822–828.
- Walker, S. (2015) Indirect gradient analysis by markov-chain monte carlo. *Plant Ecology*, **216**, 697–708.
- Walker, S.C. & Jackson, D.A. (2011) Random-effects ordination: describing and predicting multivariate correlations and co-occurrences. *Ecological Monographs*, **81**, 635–663.
- Warton, D.I., Foster, S.D., Death, G., Stoklosa, J. & Dunstan, P.K. (2014) Model-based thinking for community ecology. *Plant Ecology*, **216**, 669–682.
- Warton, D.I., Wright, S.T. & Wang, Y. (2012) Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution*, **3**, 89–101.
- Wood, S. (2012) *mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation*.

Received 1 February 2015; accepted 18 May 2015

Handling Editor: Jason Fridley

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Positive pairwise species correlations to the environment: single-predictor LVMs.

**Figure S2.** Temporal trends in species abundance.

**Table S1.** Environmental and residual correlations from individual cluster models.

**Appendix S1.** Primer on modelling species co-occurrence using latent variable models (LVM).

**Appendix S2.** fitLVM-auxiliaryfunctions.R