Probabilistic and spatially variable niches inferred from demography

Jeffrey M. Diez[1]*, Itamar Giladi[2], Robert Warren[3] and H. Ronald Pulliam[4]

[1]*Department of Botany and Plant Sciences, University of California, Riverside, CA 92521, USA;* [2]*Mitrani Department of Desert Ecology, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion 84990, Israel;* [3]*Department of Biology, SUNY Buffalo State, Buffalo, NY 14222, USA; and* [4]*Odum School of Ecology, University of Georgia, Athens, GA 30602, USA*

Summary

1. Mismatches between species distributions and habitat suitability are predicted by niche theory and have important implications for forecasting how species may respond to environmental changes. Quantifying these mismatches is challenging, however, due to the high dimensionality of species niches and the large spatial and temporal variability in population dynamics.

2. Here, we explore how probabilistic assessments of habitat suitability based on demographic models may be used to better bridge niche theory and population dynamics. We use integral projection models (IPMs) to predict population growth rates for a terrestrial orchid in response to environmental variables. By parameterizing these IPMs with hierarchical models, we develop a spatially variable measure of a species' demographic niche, which can then be compared against its distribution to test ideas about what factors control a species' distribution.

3. We found that demographic suitability of sites was not well correlated with the orchid's distribution at local scales, with many absences from microsites of high predicted suitability and occurrences in sites with low predicted suitability. However, at the population scale, abundance was positively correlated with demographic suitability of the sites. These results are consistent with dispersal limitation and source–sink dynamics at small scales but stronger distribution-suitability matching at larger landscape scales.

4. *Synthesis.* The relationships between species distributions and demographic performance underlie basic niche theory and have important implications for predicting responses to a changing environment. The complexities of these relationships will require approaches that can encapsulate what we know in probabilistic terms and allow for spatially varying niche relationships.

Key-words: Bayesian, demography, hierarchical, integral projection models, niche theory, Orchidaceae, plant–climate interactions, scale

Introduction

A species' distribution does not necessarily correspond with the distribution of its suitable habitat. These 'mismatches' are typically viewed as differences between a species' fundamental and realized niches, with an early and sustained research focus on how competition may exclude a species from part of its fundamental niche (Hutchinson 1958). Niche theory and empirical work over the last few decades have expanded this view by describing how mismatches may arise from a variety of mechanisms, including spatial and temporal variation in habitat suitability, dispersal and biotic interactions (Pulliam 2000; Fig. 1). These mismatches have practical implications, frustrating efforts to forecast species responses to climate change (Jeltsch *et al.* 2008) and the spread of invasive species (Albright *et al.* 2010). However, despite the firm theoretical

basis for mismatches, quantifying these relationships remains a significant challenge.

Part of the challenge to quantifying mismatches arises from a gulf between the neat conceptual and mathematical constructs of niche theory and the inherent variability of field studies of population dynamics. Niche theory describes how population growth or fitness of a species changes across environmental gradients (MacArthur 1972; Tilman 1980), or with species interactions (Chesson 2000). Conceptual and mathematical niche models have remained largely deterministic (Chase & Leibold 2003; but see Tilman 2004), and controlled experiments have long predominated as tests of this theory (Birch 1953). On the other hand, field studies of population dynamics are typically observational and focus on explaining the considerable temporal and spatial variation in growth rates. However, in the shadow of debates over the importance of density dependence, relationships between population dynamics and environmental variables have rarely been quantified (Sibly & Hone 2002).

*Correspondence author. E-mail: jeffdiez@gmail.com
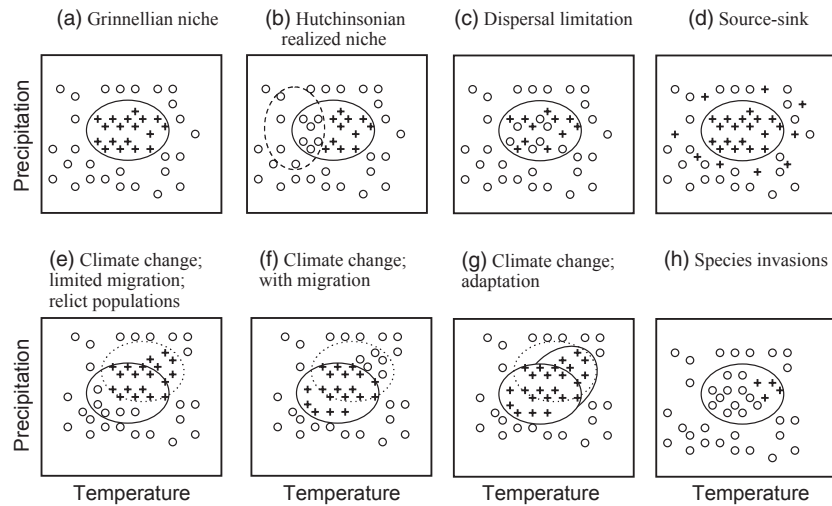
Fig. 1. Conceptual relationships between species distributions and habitat suitability (adapted from Pulliam 2000). Solid ovals demarcate the suitable habitat for a species (those environmental conditions that allow persistence as measured by positive population growth rates). The '+' symbols represent sites where the species is present, and 'o' are sites where species is absent. The top row shows classical expectations from niche theory of how a species' distribution may be related to its niche. In (a), the species is present in all suitable sites but not in unsuitable sites. In (b), the species is displaced from some suitable sites by the presence of a competitor represented by the dashed oval. Under dispersal limitation (c), the species does not arrive in all suitable sites, and source–sink dynamics (d) suggest the possibility that substantial number of individuals can occur in unsuitable habitat due to dispersal from suitable source habitats. In the bottom row (e–h), this framework is applied to climate change and invasive species scenarios. In (e–g), the originally occupied sites became warmer and wetter (shifted up and right), but the fundamental requirements for the species (solid oval) remained the same. The dotted ovals envelope current environmental conditions in previously suitable sites. A species' ability to migrate will determine whether it is absent from newly suitable sites (e) or is able to shift its distribution to match current conditions (f). Species may also adapt to the new conditions, resulting in expanded niche space (g). The same framework can represent how non-native species distributions may not match the distribution of suitable habitat, for example, because they have not had time to spread (h).

Using population dynamics to infer niche relationships is further complicated by challenges associated with spatial scale. Regional population fluctuations can sometimes be linked to climatic variation (Leirs *et al.* 1997; Coulson *et al.* 2001; Stenseth *et al.* 2002), but at broader scales than is useful for quantifying the niche. At landscape scales, many demographic field studies have documented significant spatial and temporal variation in population growth rates (Horvitz & Schemske 1995; Freville *et al.* 2004; Koop & Horvitz 2005; Jongejans *et al.* 2010), but only rarely quantified how environmental drivers affect population growth rates (but see Altwegg *et al.* 2005; Dahlgren & Ehrlén 2009). Niche relationships may also differ among populations because different abiotic and biotic contexts may alter species' responses to niche axes, and local adaptation can change the underlying niche relationships. This population differentiation is one of the major limitations to using correlations with climatic variables to predict species range shifts (Thuiller *et al.* 2008). Finally, at very local scales, individuals vary demographically due to a combination of genetic variability, environmental heterogeneity and biotic interactions (Hartgerink & Bazzaz 1984; Clark, LaDeau & Ibanez 2004; Uriarte *et al.* 2004). Incorporating this fine-scale variability, instead of averaging over it, makes predictions of species' responses to environmental drivers less certain (Clark 2003; Ibáñez *et al.* 2007) and will therefore influence the perception of niche relationships.

Thus, the substantial variability of natural populations remains challenging to explain and does not fit neatly into the conceptual or mathematical frameworks of niche theory. In this study, we build a bridge between the variability of field population dynamics and the conceptual framework of niche theory in order to quantify mismatches between species distributions and suitability (Fig. 2). As a case study of this approach, we use integral projection models (IPMs) to quantify the probability that an orchid native to Eastern North America, *Goodyera pubescens*, exhibits positive population growth in response to key environmental variables (conceptually described in Fig. 2a–b). We then use these probabilities to quantify the relationship between observed distributions and predicted performance (Fig. 2c–d). By parameterizing these IPMs using hierarchical models, we incorporate both fine-scale variability and the potential for spatially varying niches. The result is a novel perspective on the intersection of population dynamics and niche theory that respects the inherent variability in species responses to the environment but retains the framework of niche theory to explore processes underlying species distributions.

Materials and methods

STUDY DESIGN AND SPECIES

We collected demographic and environmental data in a nested design in order to capture microsite-level variation of the forest floor up to variation among populations within a watershed. The watershed was the Coweeta Hydrologic Laboratory (35° 03′ N latitude, 750 – 1500 m elevation), an LTER research site in Macon County, North
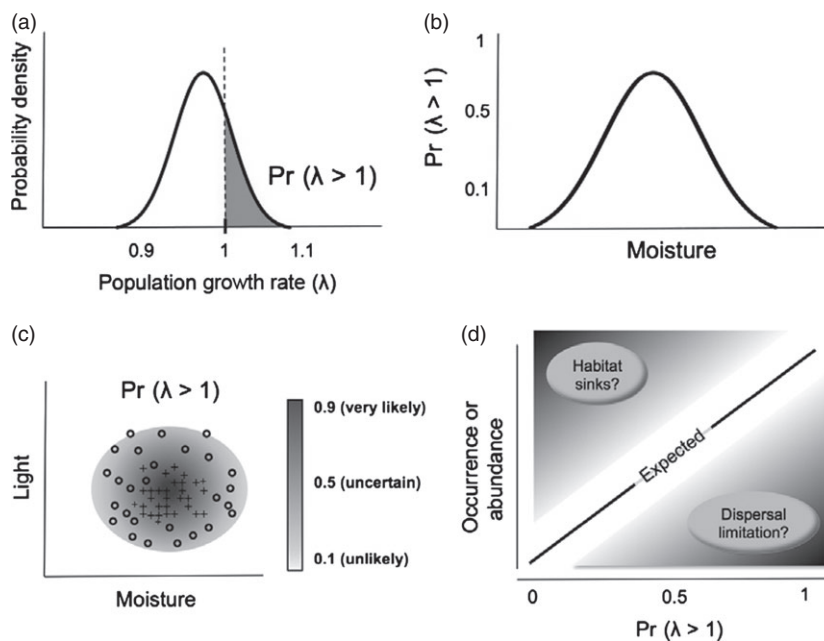
Fig. 2. Conceptual framework for a probabilistic niche. (a) The probabilities that population growth rates are >1, $Pr(\lambda > 1)$, can be estimated from posterior distributions of λ's calculated from the Bayesian demographic models. These probabilities can be calculated as a function of environmental niche axes, resulting in distributions describing the probability of obtaining positive population growth rates conditional on single (b) or multiple (c) environmental axes. The relationship between species distributions (occurrence and abundance) and population growth rates can then be modelled for particular sites by predicting the population growth rates conditional on the environment at these sites. In (C), the '+' represent sites where the species is present, and 'o' are sites where it is absent (as in Fig. 1). Predicted suitabilities can then be compared to observed patterns of occurrence and abundance (d), with the expectation (often the assumption in climate envelope modelling) that the likelihood of species' occurrence and abundance increases as the suitability of the environment for population growth increases.

Carolina. Six study grids of 480 m² each were established in mature (80+ years old) deciduous forest. These sites were chosen to reflect a wide range of elevation, temperature and precipitation conditions experienced by the species under study. To capture microsite variation, the study plots were divided into 4 m² cells, and this was the resolution for abiotic and demographic monitoring. Thus, three spatial scales will be discussed from here forward: the microsite scale (2 m × 2 m cells), population scale (480 m² grids) and overall landscape scale (reflecting that the study grids were chosen to represent a wide variety of the environments in which *G. pubescens* is found in the region).

PLANT AND ABIOTIC MONITORING

The study species, *G. pubescens*, is a perennial, evergreen, clonal orchid distributed throughout Eastern North America. Like other orchids, *G. pubescens* is dependent on colonization of seeds by the appropriate mycorrhizal fungi for germination, and it forms a symbiotic protocorm in the first year that may develop into a photosynthetic seedling. All individuals within the six study populations were individually marked and monitored for growth, reproduction and survival for a period of 6 years (1998–2004). Each of the 1848 individuals, present in 227 of the 720 grid cells, was marked with a unique flag and revisited each autumn to measure growth and reproduction. New individuals were identified each year either as seedlings or clonal offspring (through obvious breakage of above-ground rhizome connections). Thus, marked individuals were not connected via rhizome, but could be genetically identical, so our model is based on ramets and not on genets.

Soil moisture and light availability are particularly important abiotic variables for understorey wildflowers in this region and have

been shown to correlate with this species' distribution (Diez & Pulliam 2007; Fig. 3). Time domain reflectometry (TDR) was used to characterize patterns in soil moisture content in the top 12 cm of soil at 80 points within each of the study grids. Ordinary kriging using exponential variograms, accounting for directional trends in the spatial autocorrelation, was used to obtain estimates of soil moisture in all cells of the study grids (Jackson & Caldwell 1993). These measurements were taken four times a year for 3 years. Because the spatial patterns were highly correlated among time periods, average summer values were used to characterize the spatial structure of soil moisture. While these TDR readings cannot provide an absolute measure of soil moisture, the high correlation among seasons and years suggested that this was a good measure of relative soil moisture among sites.

Understorey light availability was measured as the photosynthetic photon flux densities (% PPFD) at 1 m using a hand-held AccuPAR ceptometer. Eighty ceptometer readings were taken on each grid, with concurrent measurements on a second quantum sensor taken in nearby clearings, allowing calculation of per cent transmittance of incident canopy PAR to the forest floor. Geostatistical methods similar to those used for soil moisture were used to obtain estimates of light availability for each cell. Because *G. pubescens* is an evergreen species in deciduous forests, winter light can be particularly important (Diez & Pulliam 2007), so we used this variable in analyses.

DEMOGRAPHIC MODELLING

Overall approach

To quantify demographic variability and responses to the environment, we used IPMs parameterized by mixed models. IPMs are a useful method for describing discrete-time and continuous size
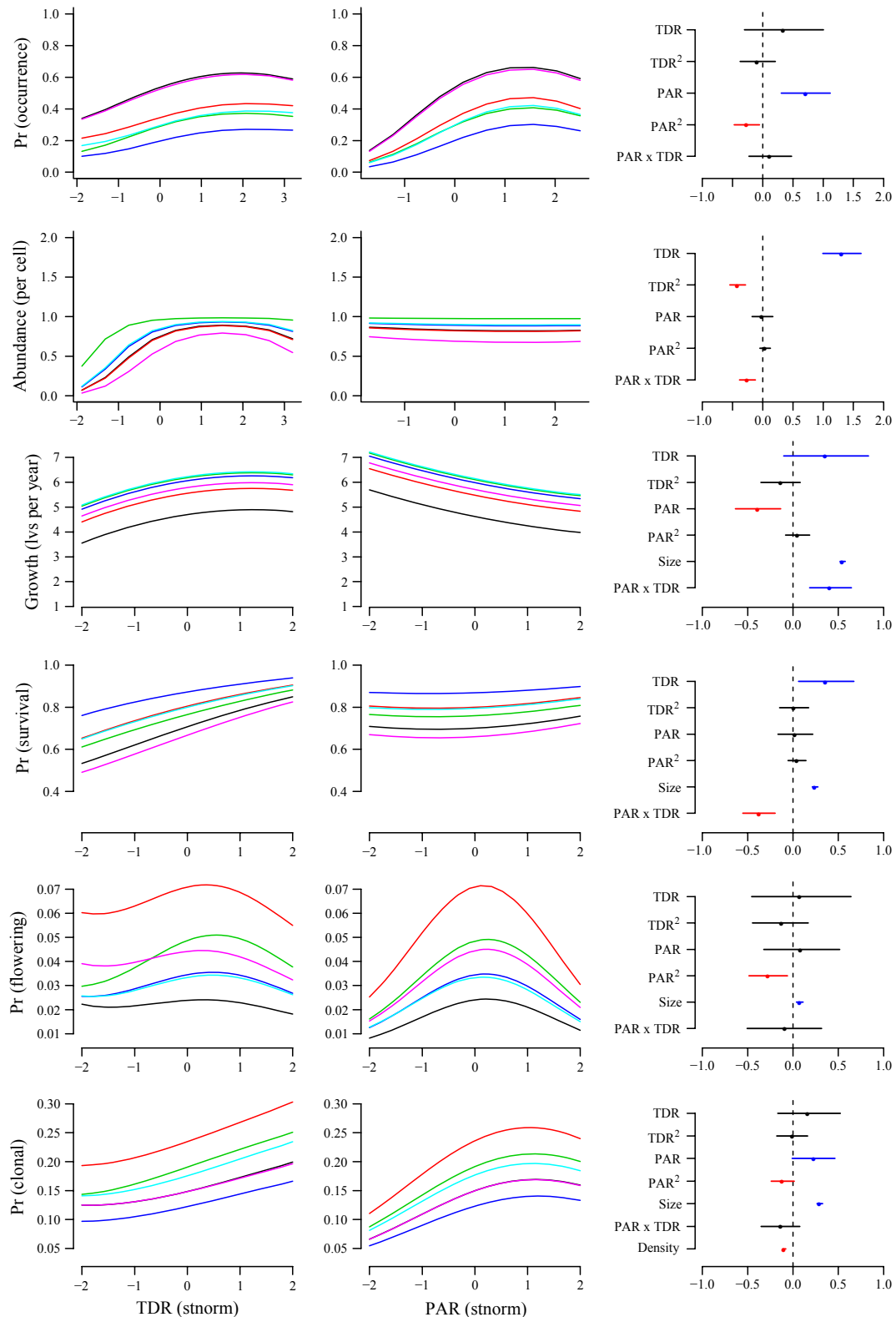
Fig. 3. Relationships between distributions, vital rates and soil moisture [time domain reflectometry (TDR)] and light availability (PAR). Left two columns: The six lines in each graph represent population-specific estimates of these relationships. These lines are the mean responses calculated within the mixed models. Responses to TDR are shown at mean PAR levels and vice versa. TDR and PAR are plotted in standard normal units, and the probability of clonal reproduction is plotted on a logit scale. The parameters describing these curves and their significance are plotted in the right column. Right column: regression coefficients from the regressions upon which the predictive lines on the left are based. The response variable being shown in each graph corresponds to the leftmost axis (i.e. from top to bottom: occurrence, abundance, growth, survival, flowering and clonality). Intervals represent 95% credible intervals of the posterior probability distributions. Parameters with at least a 90% chance of being different from zero are coloured red (if negative) and blue (if positive).

population growth (Easterling, Ellner & Dixon 2000; Metcalf *et al.* 2009). Although the general goal of the study could be accomplished with any demographic method yielding measures of population growth rates, we chose IPMs because *G. pubescens* continuously adds leaves as adults, rather than exhibiting discrete size stages (suitable for matrix models). Furthermore, IPMs are easily extended to model how environmental variables affect vital rates and population growth using mixed models (Dahlgren & Ehrlén 2009; Dalgleish *et al.* 2011). As described below, we extended previous approaches by modelling the parameters of these mixed models hierarchically, which allowed us to explore demographic responses to the abiotic environment at different scales and to test for spatially variable niches. Estimates of demographic responses to the environment from the hierarchical models can then be used to predict population growth rates using demographic models (described below) and compared to the species' distribution to assess mismatches. Although population growth rates are an imperfect measure of species niches, they are arguably the best integrative measure of plant population performance (Crone *et al.* 2011) and therefore serve as a useful tool for exploring niche relationships.

IPMs

Several good introductions to IPMs now exist in the ecological literature, and these can be consulted for background information about the technique (Ellner & Rees 2006; Williams & Crone 2006; Rees & Ellner 2009). The basic idea is that IPMs represent populations as a probability density function, $n(x, t)$, that characterizes the probability of individuals at time t being of size x. The population size distribution at time $t + 1$ is described by the kernel:

$$n(y, t + 1) = \int_L^U k(y, x, \theta) n(x, t) dx, \qquad \text{eqn 1}$$

where $k(y, x, \theta)$ describes all possible transitions between individuals of size x at time t to size y at time $t + 1$, with environmental covariates, θ, and constrained between lower (L) and upper (U) size limits. This kernel is made up of growth-survival and fecundity functions:

$$\int_L^U k(y, x, \theta) = \int_L^U [p(y, x, \theta) + f(y, x, \theta)]. \qquad \text{eqn 2}$$

The function $p(y, x, \theta)$ describes size-dependent survival and growth, and $f(y, x, \theta)$ describes size-dependent fecundity. As with matrix models, the growth-survival function allows individuals to grow, stay the same size or regress in size. When plants in the model reach the specified upper (U) or lower (L) size limits, they can become unintentionally 'evicted' from the models, artificially reducing population growth rates (Williams, Miller & Ellner 2012). In order to prevent the loss of individuals from the models, we used truncated normal distributions for the growth function (Williams, Miller & Ellner 2012), limiting the maximum size of plants to the maximum observed (39 leaves). Although this is an assumption that puts a size cap on the demographic rates, there is a very large upper tail to the observed size distribution in the field, and very few plants will ever reach this size and are unlikely to differ substantially from 39 leaf plants in their demographic rates.

These functions describing probabilities of survival, growth and fecundity were parameterized using mixed models that included plant size and normalized environmental covariates (soil moisture and light) at the 2 m × 2 m cell level. For each vital rate, two models with different hierarchical structure were fit in order to test for the appropriate scale at which demographic responses to the environment vary. In the first model, intercepts were allowed to vary among populations, but

relationships with moisture and light were assumed consistent within the landscape. Thus, the growth of individual plant i in cell c, population p and year y was modelled as:

$$\begin{aligned} \text{growth}_i = {} & \alpha_p + \delta_{\text{size}} x_i + \beta_1 M_c + \beta_2 M_c^2 + \beta_3 L_c + \beta_4 L_c^2 \\ & + \beta_5 M_c L_c + \text{RE}_y + \epsilon_i, \end{aligned} \qquad \text{eqn 3}$$

where M_c and L_c are the soil moisture and light availability, respectively, in the grid cell. The mean growth rate was allowed to vary linearly (δ_{size}) with plant size, x_i. The parameter α_p is a population-level intercept, and the β's are regression coefficients describing the effects of soil moisture, light and their interaction, on growth rates. The second-order abiotic terms (M^2 and L^2) were included in order to allow for nonlinear relationships with moisture and light (e.g. unimodal or saturating relationships). The inclusion of second-order terms was based on analyses of the species' distribution (Diez & Pulliam 2007; Fig. 3), as well as a common biological assumption that species have optimal environmental conditions, with performance declining as conditions depart from the optimum. Interactions between each environmental variable and plant size were also included to test for size-dependent responses to the environment (Williams & Crone 2006). The population-level intercepts were modelled hierarchically, where the population relationships are drawn from an overall landscape-level coefficient: for example, $\alpha_p \sim \text{Normal}(\alpha_{\text{landscape}}, \sigma^2)$. A random effect for year, RE_y, was included to account for annual variation in growth. The errors for the growth model, ϵ_i, were modelled as Gaussian.

Survival (0/1) was modelled similarly to growth, but using a generalized linear model with a *Bernoulli* sampling distribution and logit link function in order to estimate the probability of survival as a function of the same set of covariates. Clonality was estimated in a similar manner to survival, with the probability of a clonal recruit emerging in a given year being estimated as a function of ramet size and the abiotic environment. These probabilities of producing clonal offspring were added to the fecundity distribution, thus augmenting the production of new individuals each year. Fecundity from production of seeds was difficult to directly observe, so it was modelled using several pieces of information. The true numbers of seeds and protocorms (pre-seedling symbiotic stage) were unobservable in the field because the seeds are minute, and the protocorms are generally buried within leaf litter. Therefore, fecundity was modelled as the product:

$$\begin{aligned} \frac{\text{Recruits}}{\text{Individual}} = {} & Pr(\text{Flower}) \\ & \times \frac{\text{seed capsules}}{\text{flower}} \times \frac{\text{seeds}}{\text{capsule}} \times \frac{\text{protocorms}}{\text{seed}} \times \frac{\text{recruits}}{\text{protocorm}}. \end{aligned} \qquad \text{eqn 4}$$

The number of capsules per flower was directly observable, whereas the number of seeds per capsule was estimated by capsule dissections, and the number of protocorms per seed was estimated from seed packet introduction experiments (Diez 2007). The probability of flowering was modelled at the individual level, conditional on plant size, light and moisture, using logistic regression as described for survival. The other reproduction parameters were estimated independent of the environment at the grid level.

The second model structure fits to each vital rate allowed the slopes of these regressions to vary among populations instead of assuming the same relationship for all populations. The same models as above were fit to the data, but the regression coefficients, β, were allowed to vary among populations and be linked via an overall landscape-level coefficient. For example, $\beta_{1,p} \sim \text{Normal}(\beta_{1,\text{land}}, \sigma^2)$ described the population-specific responses to soil moisture, and other coefficients $\beta_{2,p}, \beta_{3,p}, \beta_{4,p}, \beta_{5,p}$ were modelled similarly. The σ^2 parameters describe the variance among populations in the species'

responses to each environmental variable, and $\beta_{1,land}$ is a parameter estimated from the data that describes the expected relationship with soil moisture for an average population in the landscape. These models also had intercepts, α_p, that varied among populations and were modelled hierarchically from an overall landscape-level intercept.

The differences between these two model structures are important to understand in terms of their corresponding ecological hypotheses. The first model with variable intercept α_p and constant β across populations assumes that the species' underlying relationship to the abiotic environment is the same in the different populations within this landscape. However, the variable intercepts shift the resulting relationships up or down in magnitude due to other, unidentified differences among the populations. The second model, which includes intercepts α_p and slopes β_p that both vary among populations, hypothesizes that the underlying relationships with the abiotic environment are different across populations. That is, a plant's growth rate, for example, may respond strongly to increasing soil moisture in one population but only weakly in another. These differences may arise either through differences in abiotic and biotic context or through local adaptation. Both would give rise to apparent differences in the species' niche relationships across populations. These two model structures are displayed visually in the appendix (Fig. S1.5 in Supporting Information).

All models were fit using Markov chain Monte Carlo (MCMC) algorithms in OpenBUGS (Lunn *et al.* 2009), as called from R (R Development Core Team 2008) using package R2OpenBUGS (Sturtz, Ligges & Gelman 2005). All parameters were given non-informative priors. Regression coefficients and overall intercepts were given Normal priors with mean 0 and variance 1000, and variance parameters were given Uniform priors between 0 and 100 on their standard deviations. Three independent chains were run for 10 000 iterations after discarding a 2000 iteration burn-in period, and convergence was assessed visually and using the Gelman–Rubin convergence statistic. Parameters were considered statistically significant if their 95% confidence interval was not overlapping zero. After convergence, the 10 000 iterations were thinned to 1000 MCMC estimates of each vital rate (growth, fecundity, survivorship, clonality) in order to parameterize an integral projection model (IPM) and calculate finite population growth rates (λ) using R code modified from Ellner & Rees (2006). We conducted model selection to explore subsets of the full model that include as covariates light, moisture and the interaction between light and moisture.

Density dependence was explored by plotting individual vital rates as a function of the number of plants within 2 m × 2 m cells. We found no evidence for the effects of density on vital rates except for clonality, which was higher at low density (Fig. S1.3). Therefore, the mixed models for clonality included a density covariate, and subsequent predictions as a function of abiotic covariates were made conditional on zero density. This approach assumes that responses to the abiotic environment at low density are the most appropriate characterization of the abiotic niche of a species.

The coefficients describing responses to the abiotic environment (β's in above equations) were used to predict population growth rates across gradients of moisture and light. This approach is similar to that used in Dahlgren and Ehrlén (2009), who explored how λ varied with soil potassium and seed predation. Here, instead of bootstrapping to obtain uncertainty of these predictions, we used MCMC parameter estimates from the mixed models to propagate all uncertainty in parameters to uncertainty of predicted growth rates. The resulting posterior probability estimates of λ were used to calculate the probability of positive population growth rates, $Pr(\lambda > 1)$, conditional on the abiotic variables, by calculating the proportion of the MCMC iterations that

yielded a $\lambda > 1$. Although a variety of other demographic measures can be calculated from IPMs, we focus on λ here because it is the most relevant for assessing the relative strength of sources and sinks.

The predicted population growth for each cell, calculated based on its measured abiotic variables, was then compared to observed presence/absence and abundance of *G. pubescens* in that cell. Generalized linear models, with Bernoulli error distribution and logit link function, were used to test the relationship between presence/absence and Pr ($\lambda > 1$). Generalized linear models with Poisson error distribution and log-link function were used to test the relationship between abundance in a cell and its predicted $Pr(\lambda > 1)$. We used an exponential regression for the relationship between abundance and the $Pr(\lambda > 1)$ at the population scale. With only six populations, there is limited ability to discern the shapes of nonlinear functions, but it is plausible that abundance may increase rapidly as habitat suitability increases. These models were fit in a Bayesian framework, and parameters significance was based on the posterior distributions of the regression coefficients.

Results

The models with fixed relationships between vital rates and abiotic variables across the landscape performed better than models with population-specific relationships. This was evident from the higher degree of significance of regression coefficients in the 'common slope' models (Fig. 3) vs. the population-specific estimates (Fig. S1.6). This result is consistent with the observed spatial structure of the abiotic environment, as soil moisture and light availability varied more among populations than within (Fig. S1.1). Therefore, all subsequent results and inference are based on the models with common slopes for all populations but varying intercepts.

MIXED MODELS

Goodyera pubescens was more likely to be present in sites with an intermediate amount of winter light (Fig. 3). Within microsites where it occurs, abundance increased with soil moisture, reaching a plateau at higher moisture (Fig. 3). Vital rates responded to soil moisture and winter light in similar ways, but with some variation among vital rates (Fig. 3). Soil moisture had positive and saturating effects on growth and survival, and no significant effects on clonality or the probability of flowering (Fig. 3). By contrast, light availability was associated with decreased growth rates and had positive but unimodal effects on flowering and clonality. All vital rates increased with plant size (Fig. 3). Significant year random effects also suggested that vital rates varied among years. Conspecific neighbourhood density had a negative effect on the probability of clonal growth (Fig. 3).

Plant size × environment interactions were not significant for survival, flowering or clonality. For growth rates, there was a significant negative interaction between light and plant size, and a significant positive interaction between moisture and size (Supporting Information). However, when these interaction terms were included, all the main effects became non-significant, and the interaction terms were significant in the same direction as the main effects (light was negative and

moisture positive). These interactions also did not change the predicted responses of vital rates across abiotic gradients (Supporting Information). Therefore, mixed models without size x environment interactions were used in IPMs as discussed below. Furthermore, model selection exploring subsets of the full model including light, moisture and their interaction showed that for most vital rates, the full model was most supported (Appendix S2). Clonality was the one rate for which the model without interactions was better supported by the deviance information criteria (DIC) than the full model. However, inclusion of the interaction term had negligible effect on the outcomes (Figs S2.1 and S2.2), and so the interaction was included for all vital rates for consistency. The IPMs were therefore parameterized using the full models that included effects of light, moisture and the interaction between light and moisture on each vital rate.

IPMS, POPULATION GROWTH RATES

Based on the effects of abiotic variables on vital rates and population-specific intercepts in the models, the predicted population growth rates varied among populations but exhibited a similar pattern (Fig. 4). Each population had a maximum probability of persistence at higher soil moisture and lower winter light levels. Each population also had areas of environmental space with a wide range of predicted probabilities of positive population growth. The occurrence and absence of the species from cells within each population are plotted as '+' and 'o', respectively. It is difficult to assess the pattern of mismatch from these raw graphs, but the predicted probability of persistence at each of these points form the basis for the regressions shown in Fig. 5.

MISMATCHES

The probability of occurrence and abundance of *Goodyera* were not strongly predicted by the probability of positive population growth at the scale of 2 m × 2 m microsites (Fig. 5a, b). In fact, the probabilities of occurrence and abundance were remarkably flat in relation to predicted suitability. In contrast to this lack of relationships at small scales, the abundance of populations was positively related to the probability of positive population growth (Fig. 5c). The nonlinear, exponential shape of the relationship suggested that sites with high predicted suitability could have a range of abundances.

Discussion

Hutchinson defined the niche as the set of environmental conditions under which a species can persist (Hutchinson 1958). Persistence entails positive population growth rates when rare, so quantifying a species' niche requires measuring population growth rates in relation to environmental niche axes. This proves remarkably challenging and has been seldom done, due in part to the high variability of species' demography in natural systems. In this study, we have explored how a probabilistic approach (both conceptual framework and quantitative

analyses) can help express population dynamics in terms of niche theory while incorporating the high natural variability in species' performance and difficulties of discerning complex responses to niche axes in the field.

MISMATCHES: MECHANISMS AND IMPLICATIONS

We found that the match between demographic suitability and the species' distribution depended on spatial scale. At a small, microsite scale (4 m^2), the occurrence and abundance of *G. pubescens* were not well correlated with predicted population growth rates, while its abundance at the scale of populations was positively correlated with predicted population growth rates (Fig. 5). This difference between scales is useful for discerning the mechanisms underlying species distributions. Mismatches between species distributions and habitat suitability may arise from a variety of mechanisms, including biotic interactions (e.g. competition), dispersal and changes in suitability over time (Fig. 1). The many unoccupied microsites with high probabilities for self-sustaining population growth rates are consistent with dispersal limitation, and the occupied sites with low predicted population growth rates are consistent with source–sink theory (Pulliam 2000). Together, these results suggest that occupancy at local scales is highly stochastic (Diez & Giladi 2011). Many chance events may lead to presence or absence in a site independent of its suitability. Demographic stochasticity is also particularly relevant at small scales and may cause local extinctions (Hanski 1998).

In contrast to these relationships within microsites, there was a positive relationship between abundance and the probability of positive population growth rate among populations (Fig. 5c). This suggests that the demographic suitability of different parts of a landscape will indeed result in differences in population sizes. These relationships are scale dependent, however, and it is certainly possible to have large suitable areas with no individuals or low abundances simply because the species has not yet arrived (Pinto & MacDougall 2010). Such a scenario is particularly likely under rapid climate change or species introductions to a novel landscape (Fig. 1).

At any given scale, the degree to which species exhibit mismatches is likely to depend on various life-history characteristics (Moore & Elmendorf 2006). In this study, *G. pubescens* relies on mycorrhizal relationships with a specialized group of soil fungi (genus *Tulasnella*) to reach photosynthetic stage (McCormick, Whigham & O'Neill 2004), so recruitment patterns can be influenced by small-scale soil heterogeneity that affects fungal distributions (Diez 2007). Seeds co-introduced to field sites with suitable fungi are more likely to develop into protocorms than those without (J. M. Diez, unpubl. data), but very little is known about the distribution of these fungi. Thus, the environmental conditions conducive to fungal growth may thus constitute a 'hidden' niche dimension for the orchid. Also, despite prolific seed production when flowers are produced, the low frequency of flowering may cause seed and dispersal limitation in some areas, and clonal reproduction is a relatively slow spatial process. For many

(a) Overall niche


(b) Population level niches
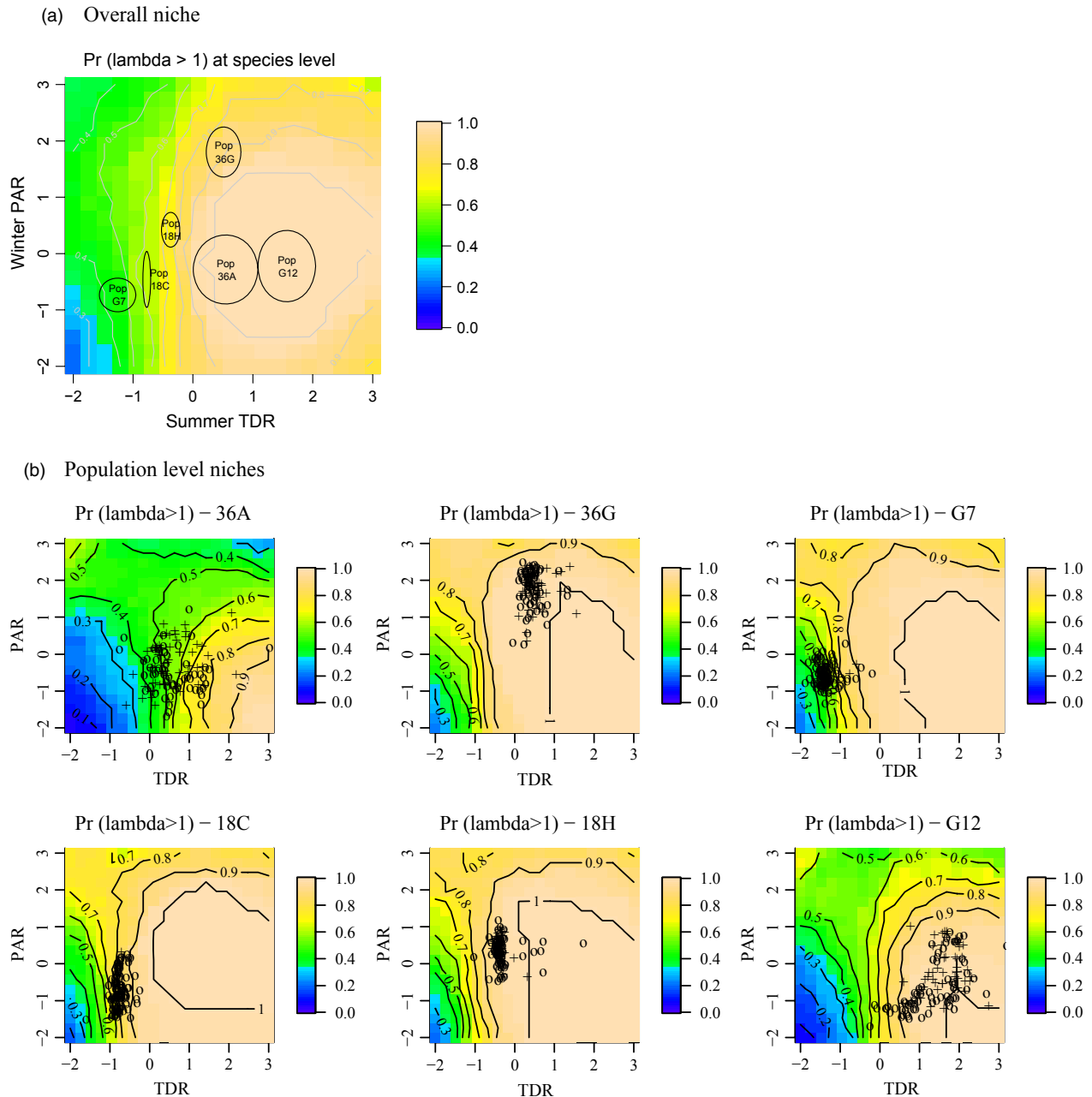
Fig. 4. Correspondence between habitat suitability and species distributions at different scales. Contours show predicted probabilities of self-sustaining populations ($\lambda > 1$) conditional on soil moisture [time domain reflectometry (TDR)] and winter light (PAR). (a) The abundance of *Goodyera pubescens* in each population (written as '$N = ...$') is plotted over the landscape-level predicted probabilities of positive population growth rates. Intervals show the standard deviation of observed TDR and PAR values in each population. (b) The presence (+) or absence (o) of *G. pubescens* in 2 m × 2 m cells is plotted on population-level estimates of the probabilities of positive population growth. The estimates of each populations' growth rates are different depending on whether one considers the overall niche or the population-specific niche, due to different estimates in each case of the parameters describing the niche relationships.

species, occurrence in unsuitable sites or marginal habitat may also result from passive dispersal from nearby suitable sites (Kadmon & Tielborger 1999). Species with small, wind-dispersed seeds such as *G. pubescens* can access a relatively large spatial extent (Singleton *et al.* 2001), which may include a great deal of marginal habitat.

Temporal changes in environmental conditions may also be a common mechanism leading to mismatches. Habitat suitability may change over time due to local disturbances such as treefalls, or larger-scale processes such as forest succession or systematic changes in climate. Under systematic climate change, the number of 'relict populations' in unsuitable habitat is expected to increase at the trailing edge, as well as the amount of unoccupied but suitable habitat on the leading edge. These expectations lead to predictable patterns of mismatch under climate change scenarios (Fig. 1). Studies
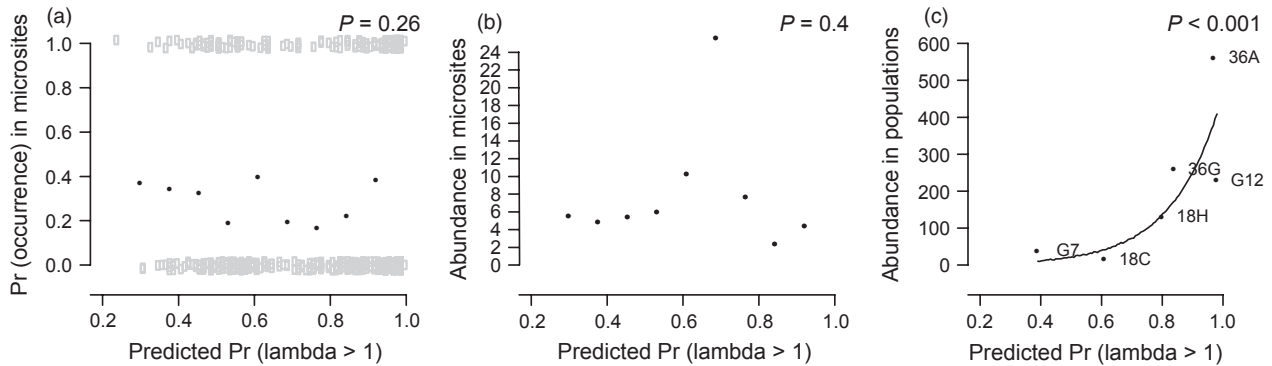
Fig. 5. The relationships between *Goodyera pubescens* distribution and the probability that $\lambda > 1$ at different scales. The relationship between the probability that $\lambda > 1$ and the probability of occurrence (a) within microsites (2 m × 2 m grid cells), abundance within grid cells (b) and the abundance within populations (c). The grey circles in (a) are raw data, and the solid points are proportions of occupied cells within nine evenly spaced bins of predicted lambda values, for easier visualization of the pattern. Similarly, the points in (b) are average abundances within nine evenly spaced bins. Significance of these relationships was assessed using the posterior distributions of the regression coefficients. The *P* value given is analogous to a null-hypothesis test, calculated as the probability that the slope of the line is less than zero (for slopes with positive mean estimates) or greater than zero (for slopes with negative mean estimates).

linking demography to abiotic variables should be conducted across a wider portion of a species' range in order to test for this signature of climate change (Doak & Morris 2010).

In addition to the above biological reasons for mismatches between distribution and demographic suitability, it is also possible that unmeasured but important environmental variables could obscure the relationship. Such 'hidden niche dimensions' are a perennial challenge in field studies of the niche. It is certain that *G. pubescens* has significant responses to more environmental variables than light and moisture, and these other effects may contribute to the apparent unpredictability of habitat suitability in our study at fine scales. As additional variables are identified through a combination of observation and experiment, predictions of suitability can be refined and the uncertainty should decrease. Nonetheless, a more precise understanding of multiple niche axes will not assure a tight fit between predicted suitability and observed distributions, but the mismatches will be more clearly attributable to other mechanisms. This approach will therefore be particularly useful for species with relatively well-known niche axes.

SPATIALLY VARYING NICHES

Although niches are often thought of (and modelled as) a set of species-level characteristics, demographic responses to environmental variables are not homogenous within species. Differences in species' observed realized niches may result from covarying abiotic or biotic environment and from local adaptation (Pulliam and Waser 2010). Even at small scales, individuals vary genetically and experience different local neighbourhood constraints. At larger spatial scales, populations may experience different abiotic and biotic environments that affect perceived responses to niche axes. For example, demographic performance may increase with soil moisture in dry sites, but decrease in wet sites because of an associated increase in competition or pathogen pressure. These realized differences can also drive local adaptation that changes the

underlying responses of the species to environmental gradients. This population differentiation, widely studied in population genetics, is one of the major limitations to using correlations with climatic variables to predict species range shifts (Thuiller *et al.* 2008).

In this study, we show how hierarchical models may be used to explore spatially variable niches in two ways. First, underlying demographic responses to environmental variables were allowed to vary among populations. In this case study, we did not find support for variation among populations in their responses to the environment, but this may be due to the spatial scale of the study. The abiotic variability among populations was greater than the variability within populations (Fig. S1.1), and there may be enough gene flow within this landscape to minimize local adaptation. Therefore, variation in a species' underlying response to niche axes may only be expected over larger spatial scales. This approach may thus prove very useful for estimating variable responses to environmental conditions across species ranges. The second form of spatial variability was represented in the models with the population-level intercepts. These allowed different populations to have different demographic outcomes given the same underlying responses to the environment. These differences reflect the unknown factors that vary among populations and shape the performance of the species (e.g. differences in competitors), and as more niche variables are identified, these population effects would be expected to diminish.

OUTSTANDING CHALLENGES AND FUTURE DIRECTIONS

Several important extensions to this work are important for building a broader understanding of how species distributions are related to habitat suitability. First, comparative studies of multiple species will help to understand how mismatches depend on species traits. For example, high dispersal ability is expected to increase the number of propagules landing in unsuitable habitat, while restricted dispersal is expected to

keep populations close to suitable habitat (Pulliam 2000), but this prediction is largely untested. Similarly, life span may influence the degree to which individuals are able to persist in temporally changing habitats, with long-lived species potentially persisting in relict populations after suitable conditions have changed (Hampe 2005). The presence–absence of short-lived species may be a better barometer of the suitability of different sites.

Secondly, there are difficult practical and philosophical questions about temporal scale that need to be addressed with longer-term data and additional modelling tools. The relationship between population dynamics and habitat suitability may vary widely over time, and short-term population growth does not necessarily reflect long-term habitat suitability (Bierzychudek 1999). Stochastic models may be used to estimate the effect of temporal variation in the environment on growth rates when such information is available (Metcalf *et al.* 2009; Evans, Holsinger & Menges 2010), although the correspondence to niche theory (Fig. 1) then becomes less clear. Further development of density-dependent models will also be necessary to accurately assess demographic suitability over longer time-scales. In this study, the relatively young forests and longevity of the study species may limit the importance of density dependence, but for many species, population growth rates will only be a relevant measure of suitability once density is accounted for. Thus, continued efforts to tease apart environmental forcing and endogenous dynamics of population growth should help further understand niche relationships.

Finally, although observational demographic studies are critical for quantifying the performance of natural populations across environmental and biotic gradients, experimental introductions into unoccupied habitats are extremely useful for testing ideas about niche axes (Moore 2009; Warren & Bradford 2011). Experiments have the potential to clearly demonstrate absence from suitable habitat (Primack & Miao 1992) and also absence from demographically unsuitable sites (Moore & Elmendorf 2006). Integration of these experimental and demographic modelling approaches using a probabilistic framework has potential to clarify how distributions are related to habitat suitability.

Acknowledgements

References

Albright, T.P., Chen, H., Chen, L. & Guo, Q. (2010) The ecological niche and reciprocal prediction of the disjunct distribution of an invasive species: the example of *Ailanthus altissima*. *Biological Invasions*, **12**, 2413–2427.

Altwegg, R., Dummermuth, S., Anholt, B.R. & Flatt, T. (2005) Winter weather affects asp viper *Vipera aspis* population dynamics through susceptible juveniles. *Oikos*, **110**, 55–66.

Bierzychudek, P. (1999) Looking backwards: assessing the projections of a transition matrix model. *Ecological Applications*, **9**, 1278–1287.

Birch, L.C. (1953) Experimental background do the study of the distribution and abundance of insects. *Ecology*, **34**, 698–711.

Chase, J.M. & Leibold, M.A. (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. (2003) Uncertainty and variability in demography and population growth: a hierarchical approach. *Ecology*, **84**, 1370–1381.

Clark, J.S., LaDeau, S. & Ibanez, I. (2004) Fecundity of trees and the colonization-competition hypothesis. *Ecological Monographs*, **74**, 415–442.

Coulson, T., Catchpole, E.A., Albon, S.D., Morgan, B.J.T., Pemberton, J.M., Clutton-Brock, T.H., Crawley, M.J. & Grenfell, B.T. (2001) Age, sex, density, winter weather, and population crashes in Soay sheep. *Science*, **292**, 1528–1531.

Crone, E.E., Menges, E.S., Ellis, M.M., Bell, T., Bierzychudek, P., Ehrlen, J. *et al.* (2011) How do plant ecologists use matrix population models? *Ecology Letters*, **14**, 1–8.

Dahlgren, J.P. & Ehrlén, J. (2009) Linking environmental variation to population dynamics of a forest herb. *Journal of Ecology*, **97**, 666–674.

Dalgleish, H., Koons, D., Hooten, M., Moffet, C. & Adler, P. (2011) Climate influences the demography of three dominant sagebrush steppe plants. *Ecology*, **92**, 75–85.

Diez, J.M. (2007) Hierarchical patterns of symbiotic orchid germination linked to adult proximity and environmental gradients. *Journal of Ecology*, **95**, 159–170.

Diez, J.M. & Giladi, I. (2011) Scale-dependence of habitat sources and sinks. *Sources, Sinks and Sustainability* (eds J. Liu, V. Hull, A.T. Morzillo & J.A. Wiens), pp. 1–26. Cambridge University Press, Cambridge, United Kingdom.

Diez, J.M. & Pulliam, H.R. (2007) Hierarchical analysis of species distributions and abundance across environmental gradients. *Ecology*, **88**, 3144–3152.

Doak, D.F. & Morris, W.F. (2010) Demographic compensation and tipping points in climate-induced range shifts. *Nature*, **467**, 959–962.

Easterling, M.R., Ellner, S.P. & Dixon, P.M. (2000) Size-specific sensitivity: applying a new structured population model. *Ecology*, **81**, 694–708.

Ellner, S.P. & Rees, M. (2006) Integral projection models for species with complex demography. *The American Naturalist*, **167**, 410–428.

Evans, M., Holsinger, K. & Menges, E. (2010) Fire, vital rates, and population viability: a hierarchical Bayesian analysis of the endangered Florida scrub mint. *Ecological Monographs*, **80**, 627–649.

Freville, H., Colas, B., Riba, M., Caswell, H., Mignot, A., Imbert, E. & Olivieri, I. (2004) Spatial and temporal demographic variability in the endemic plant species *Centaurea corymbosa* (Asteraceae). *Ecology*, **85**, 694–703.

Hampe, A. (2005) Fecundity limits in *Frangula alnus* (Rhamnaceae) relict populations at the species' southern range margin. *Oecologia*, **143**, 377–386.

Hanski, I. (1998) Metapopulation dynamics. *Nature*, **396**, 41–49.

Hartgerink, A.P. & Bazzaz, F.A. (1984) Seedling-scale environmental heterogeneity influences individual fitness and population-structure. *Ecology*, **65**, 198–206.

Horvitz, C.C. & Schemske, D.W. (1995) Spatiotemporal variation in demographic transitions of a tropical understory herb – projection matrix analysis. *Ecological Monographs*, **65**, 155–192.

Hutchinson, G.E. (1958) Concluding remarks. *Cold Spring Harbour Symposium on Quantitative Biology*, **22**, 415–427.

Ibáñez, I., Clark, J.S., LaDeau, S. & Lambers Hille Ris, J. (2007) Exploiting temporal variability to understand tree recruitment response to climate change. *Ecological Monographs*, **77**, 163–177.

Jackson, R.B. & Caldwell, M.M. (1993) Geostatistical patterns of soil heterogeneity around individual perennial plants. *Journal of Ecology*, **81**, 683–692.

Jeltsch, F., Moloney, K.A., Schurra, F.M., Kochy, M. & Schwager, M. (2008) The state of plant population modelling in light of environmental change. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 171–189.

Jongejans, E., Jorritsma-Wienk, L., Becker, U., Dostál, P., Mildén, M. & De Kroon, H. (2010) Region versus site variation in the population dynamics of three short-lived perennials. *Journal of Ecology*, **98**, 279–289.

Kadmon, R. & Tielborger, K. (1999) Testing for source-sink population dynamics: an experimental approach exemplified with desert annuals. *Oikos*, **86**, 417–429.

Koop, A.L. & Horvitz, C.C. (2005) Projection matrix analysis of the demography of an invasive, nonnative shrub (*Ardisia elliptica*). *Ecology*, **86**, 2661–2672.

Kuss, P., Rees, M., Ægisdóttir, H.H., Stephen, P. & Stöcklin, J. (2008) Evolutionary demography of long-lived monocarpic perennials: a time-lagged integral projection model. *Journal of Ecology*, **96**, 821–832.

Leirs, H., Stenseth, N.C., Nichols, J.D., Hines, J.E., Verhagen, R. & Verheyen, W. (1997) Stochastic seasonality and nonlinear density-dependent factors regulate population size in an African rodent. *Nature*, **389**, 176–180.

Lunn, D., Spiegelhalter, D., Thomas, A. & Best, N. (2009) THE BUGS project: evolution, critique, and future directions. *Statistics in Medicine*, **28**, 3049–3067.

MacArthur, R.H. (1972) *Geographical Ecology*. Harper and Row, New York, NY.

McCormick, M.K., Whigham, D.F. & O'Neill, J. (2004) Mycorrhizal diversity in photosynthetic terrestrial orchids. *New Phytologist*, **163**, 425–438.

Metcalf, C., Horvitz, C., Tuljapurkar, S. & Clark, D. (2009) A time to grow and a time to die: a new way to analyze the dynamics of size, light, age, and death of tropical trees. *Ecology*, **90**, 2766–2778.

Moore, K. (2009) Fluctuating patch boundaries in a native annual forb: the roles of niche and dispersal limitation. *Ecology*, **90**, 378–387.

Moore, K.A. & Elmendorf, S.C. (2006) Propagule vs. niche limitation: untangling the mechanisms behind plant species' distributions. *Ecology Letters*, **9**, 797–804.

Pinto, S.M. & MacDougall, A.S. (2010) Dispersal limitation and environmental structure interact to restrict the occupation of optimal habitat. *The American Naturalist*, **175**, 675–686.

Primack, R.B. & Miao, S.L. (1992) Dispersal can limit local plant distribution. *Conservation Biology*, **6**, 513–519.

Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349–361.

Pulliam, H.R. & Waser, N.M. (2010) Ecological invariance and the search for generality. *The Ecology of Place: Contributions of Place-Based Research to Ecological Understanding* (eds I. Billick & M.V. Price), pp. 69–92. University of Chicago Press, Chicago.

R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rees, M. & Ellner, S. (2009) Integral projection models for populations in temporally varying environments. *Ecological Monographs*, **79**, 575–594.

Sibly, R.M. & Hone, J. (2002) Population growth rate and its determinants: an overview. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **357**, 1153–1170.

Singleton, R., Gardescu, S., Marks, P.L. & Geber, M.A. (2001) Forest herb colonization of postagricultural forests in central New York State, USA. *Journal of Ecology*, **89**, 325–338.

Stenseth, N.C., Mysterud, A., Ottersen, G., Hurrell, J.W., Chan, K.S. & Lima, M. (2002) Ecological effects of climate fluctuations. *Science*, **297**, 1292–1296.

Sturtz, S., Ligges, U. & Gelman, A. (2005) R2WinBUGS: a package for running WinBUGS from R. *Journal of Statistical Software*, **12**, 1–16.

Thuiller, W., Albert, C., Araujo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgely, G.F., Paterson, J., Schurr, F.M., Sykes, M.T. & Zimmermann, N.E. (2008) Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology Evolution and Systematics*, **9**, 137–152.

Tilman, D. (1980) Resources: a graphical-mechanistic approach to competition and predation. *American Naturalist*, **116**, 362–393.

Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 10854–10861.

Uriarte, M., Canham, C.D., Thompson, J. & Zimmerman, J.K. (2004) A neighborhood analysis of tree growth and survival in a hurricane-driven tropical forest. *Ecological Monographs*, **74**, 591–614.

Warren, R.J. & Bradford, M.A. (2011) The shape of things to come: woodland herb niche contraction begins during recruitment in mesic forest microhabitat.

Proceedings of the Royal Society of London. Series B, Biological Sciences, **278**, 1390–1398.

Williams, J.L. & Crone, E.E. (2006) The impact of invasive grasses on the population growth of Anemone patens, a long-lived native forb. *Ecology*, **87**, 3200–3208.

Williams, J.L., Miller, T.E.X. & Ellner, S.P. (2012) Avoiding unintentional eviction from integral projection models. *Ecology*, **93**, 2008–2014.

Supporting Information

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

Appendix S1.

Figure S1.1. Exploratory plots of abiotic conditions in the six populations.

Figure S1.2. Exploratory plots of vital rates.

Figure S1.3. Density dependence: clonal reproduction was the only vital rate with density dependence detected.

Figure S1.4. Parameter values from the models that included size x abiotic interactions.

Figure S1.5. Hierarchical study design and model structure.

Figure S1.6. Testing for differences in abiotic relationships among populations.

Appendix S2.

Table S1. Model selection.

Figure S2.1. Full model.

Figure S2.2. No interactions.

Figure S2.3. Light (PAR) only.

Figure S2.4. Soil moisture (TDR) only.

Appendix S3. Model equations with mean parameter estimates and standard deviations in parentheses.

Table S3.1. Intercept parameter values for each population and vital rate (α_p in equations above).