Assessing the accuracy of density-independent demographic models for predicting species ranges

Keywords: demographic distribution model, density-dependence, invasion risk map, matrix population model, range dynamics, range shifts, species distribution model

Abstract:

Accurately predicting species ranges is a primary goal of ecology. Demographic distribution models (DDMs), which correlate underlying vital rates (e.g. survival and reproduction) with environmental conditions, can potentially predict species ranges through time and space. However, tests of DDM accuracy across wide ranges of species' life histories are surprisingly lacking. Using simulations of 1.5 million hypothetical species' range dynamics, we evaluated when DDMs accurately predicted future ranges, to provide clear guidelines for the use of this emerging approach. We limited our study to deterministic demographic models ignoring density dependence, since these models are the most commonly used in the literature. We found that density-independent DDMs overpredicted extinction if populations were near carrying capacity in the locations where demographic data were available. However, DDMs accurately predicted species ranges if demographic data were limited to sites with mean initial abundance less than one half of carrying capacity. Additionally, the DDMs required demographic data from at least 25 sites, over a short time-interval (≤10 time-steps), as populations initially below carrying capacity can saturate in long-term studies. For species with demographic data from many low density sites, DDMs predicted occurrence more accurately than correlative species distribution models (SDMs) in locations where the species eventually persisted, but not where the species went extinct. These results were insensitive to differences in simulated dispersal, levels of environmental stochasticity, the effects of the environmental variables, and the functional forms of density dependence. Our findings suggest that deterministic, density-independent DDMs are appropriate for applications where locating all possible sites the species might occur in is prioritized over reducing false presence predictions in absent sites. This makes DDMs a

Deleted: demographic distribution models

Deleted: that ignore density dependence (1) accurately predict range dynamics for small, fast-growing populations, but (2) overpredict extinction for slow-growing populations, and (3) overpredict

Deleted: a collection of

Deleted: 2-5

Deleted: as

Deleted: locally

Deleted: in absent sites, more accurately

Deleted: 2as

Deleted: . For fast-growing populations, demographic distribution models accurately predict occurrence as long as the population is at <65% of carrying capacity in the sites where the demographic data are collected. For slow-growing populations, DDMs greatly overpredict extinction

Deleted: more

Deleted: where

Deleted: all possible future locations of a species

Deleted:

Deleted: minimizing

Deleted: a future

Deleted: identifying present sites is more important than absent sites, where demographic data from small populations are available

promising tool for mapping invasion risk. However, demographic data are often collected at sites where a species is abundant. Density-independent DDMs are inappropriate in this case

Deleted: and predicting ranges of quickly recovering threatened species

Deleted: many species are either slow-growing or their

Deleted:

Deleted: emographic distribution model

Deleted: ese settings

Introduction

Spatial projections of species occurrence and persistence are essential for developing ecological theory and improving environmental management (Guisan et al. 2013, Meyer et al. 2015, Briscoe et al. 2019). While scientists and environmental agencies commonly correlate static presence/absence data with environmental variables to project species ranges (Elith and Leathwick 2009, Guisan et al. 2013, Hijmans et al. 2017), the accuracy and utility of these projections have been criticized, especially when projecting distributions in time or to novel environments (Pearson and Dawson 2003, Thuiller et al. 2014, Zurell et al. 2016, Cabral et al. 2017, Briscoe et al. 2019). Unfortunately, due to current, rapid, environmental change, projections in time and to novel environments are urgently needed (Ackerly et al. 2010). Demographic distribution models (DDMs), which model vital rates, such as survival, development, and reproduction, as functions of environmental variables, have been proposed as a promising alternative for generating these projections (Merow et al. 2014, 2017, Zurell et al. 2016, Briscoe et al. 2019). DDMs are promising because they model the underlying mechanisms that drive occurrence (Normand et al. 2014, Cabral et al. 2017), mechanisms which may continue to hold in new environments. Yet the accuracy and limitations of DDMs are poorly understood.

The first step to building a DDM is regressing vital rates against environmental factors to determine trajectories of population size through time, which then may be projected into the future to predict species occurrence and abundance (Villellas et al. 2015, Ehrlén et al. 2016, Cabral et al. 2017, Csergő et al. 2017). This explicit focus on capturing how environmental variables on vital rates differs from other approaches using demographic models to to project species' ranges. For example, coupled niche-population models use stochastic demographic models to project range dynamics, but vital rates are typically fixed, with environmental effects

Formatted: Highlight

captured by constraining potential habitat and carrying capacity based on modelled habidtat suitability (Keith et al. 2008, Fordham et al. 2013, 2018). Such approaches allow one to explicitly project range dynamics with limited demographic data accross sites, but their reliance on occurance data to model environmental cosntraints means that they cab suffer from many of the same drawbacks as correlative SDMs (Briscoe et al. 2019). Unfortunately, DDMs also have important drawbacks. Even the simplest DDMs require population abundance data through time (Buckley et al. 2010), or detailed demographic data tracking many individuals and their offspring within a field season (Merow et al. 2014). In both cases, these data must be collected at multiple, geographically, and climatically dissimilar sites, and classified by age, and/or size of development (Caswell 2001, Needham et al. 2018). This requirement of spatial and temporal replication is challenging in its own right. Therefore, DDMs typically ignore the effect of intraspecific competition on survival or reproduction, despite tools for incorporating density-dependence effects in demographic models (Cushing et al. 2002, Dahlgren et al. 2014, Teller et al. 2016). To our knowledge, the vast majority of DDMs, parameterized with field data, to project species ranges, ignore density-dependent effects (Buckley et al. 2010, Barbraud et al. 2011, Merow et al. 2014, 2017, Sheth and Angert 2018, Needham et al. 2018) but see (Pagel et al. 2020) for an exception. Such density-independent DDMs predict occupancy by linking environmental variables to long-term population growth rate, λ, through the variables' effects on vital rates in matrix-population or integral-projection models (Buckley et al. 2010, Merow et al. 2014, 2017). The implicit logic is that if $\lambda>1$, the local population is predicted to persist; if $\lambda<1$, the population is predicted to go locally extinct in the long-term.

However, estimates of λ from density-independent demographic models do not necessarily reflect long-term persistence if the population experiences density-dependent

Formatted: Highlight

Formatted: Highlight

Deleted: such models

Formatted: Highlight

survival, development, and/or fecundity. If fecundity or survival are lower at high population densities due to, for example, competition for resources, populations can approach a long-term equilibrium abundance – which we will refer to as carrying capacity. It is likely that demographers often collect data where populations are abundant, *i.e.* close to carrying capacity (Quintana-Ascencio et al. 2018, Fournier et al. 2019). This is because large populations are easier to find and researchers often go where healthy populations are known to exist, not to fringe populations, likely to produce small datasets. However, in healthy populations near carrying capacity, there may be limited population growth, even if the site is highly suitable. Demographic models fit to data from these sites should produce $\lambda \sim 1$, and therefore, an estimate of $\lambda < 1$ could simply reflect measurement error, disturbance, or temporary declines after populations exhausted their resources. In short, if demographic data are collected in field sites where the species is abundant, a density-independent DDM using these data could erroneously map prime habitat as uninhabitable. Therefore, it is no surprise that empirical studies often find estimates of λ uncorrelated or even negatively correlated with habitat suitability or species occurrence (Diez et al. 2014, Thuiller et al. 2014, Csergő et al. 2017). In contrast, there are at least two cases of density-independent DDMs built from demographic data restricted to sites with small populations (e.g. an invasive species and species experiencing high levels disturbance). In these cases, λ successfully predicted species occurrence (Merow et al. 2014, 2017).

Given the mixed success of initial attempts to predict species ranges using density-independent demographic models, we set out to determine general guidelines for when these models can predict species occupancy accurately. We achieved this by simulating range dynamic data, and observers sampling the data to build DDMs. We then compared DDM predictions

against long term occupancy, assessed DDM accuracy, and correlated DDM accuracy with various species and population characteristics. Finally, we compared DDM predictions to predictions from standard correlative species distribution models (SDMs). In face of limited data available to validate predictions of species range dynamics, our simulated approach provides a tool for assessing DDM accuracy. Our computational framework has many advantages over traditional validation and sensitivity analyses using real-world data, including: increased repeatability, transparency, sample sizes, and control over environmental and historical factors (Zurell et al. 2010) – all of which help improve the generality of the results.

Material and methods

Our study involved three separate processes: (1) simulating population dynamics for hypothetical species using a stochastic, negative density-dependent, spatially-explicit, stage-structured population model with two life stages (juvenile, adult) and juvenile dispersal (see Fig 1 for a graphical description of the model); (2) simulating sampling by field workers conducting demographic surveys across a subset of the species' habitat; and (3) fitting demographic distribution models, computed from the sampled field data. To determine the characteristics of species that can successfully be modeled using demographic distribution models, we simulated the range dynamics of 1.5 million hypothetical species that differed in their maximum survival rates at each life stage, maximum fecundity at low densities, maximum carrying capacity, response to environmental variables, stochastic variability in survival, initial population densities, and proportion of the population that disperses at each time step. We then determined if we could draw general conclusions about the species for which DDMs made accurate *vs.* inaccurate predictions of species occurrence.

Deleted: stochastic, negative density-dependent, spatially-explicit,

Deleted: 100,000

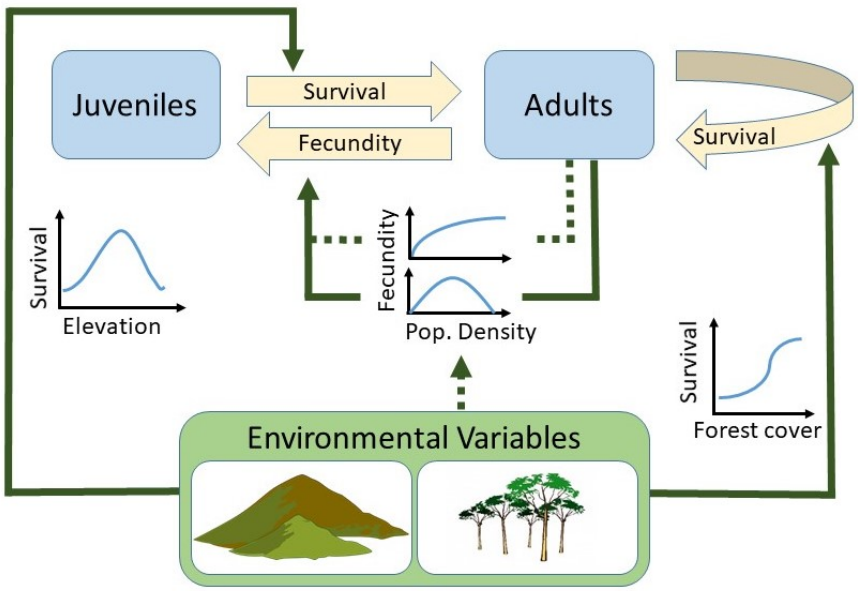


Fig 1. A graphical depiction of the simulation model. Solid dark arrows represent the effects of environmental variables and population density on vital rates (thick yellow arrows) in the baseline scenario. Dashed arrows are for effects absent in the baseline scenario, but which are tested in the sensitivity analysis. Blue curves show the assumed functional relationships between the variables. Note that both elevation and forest cover affect survival in both life stages, but repeated arrows are omitted to improve readability. For the beverton-Holt and logistic fecundity curves, the fecundity axis is total offspring (this is rescaled to expected per-capita offspring in the model description).

Simulated population dynamics - survival

We simulated population dynamics using a simple stochastic model. The probability of survival for an individual in life stage *s* and site *i*, $\varphi(i, s)$, was a function of life stage, and environmental conditions,

$$\varphi(i,s) = \frac{\psi_s e^{u(i,s)}}{1 + e^{u(i,s)}}, \qquad \text{(eqn. 1)}$$

$$u(i,s) = \sum_{j=1}^{n_e} \left(\alpha_{1,s,j} x_{i,j} + \alpha_{2,s,j} x_{i,j}^2\right) + \alpha_0 + \epsilon_{i,s,t}. \qquad \text{(eqn. 2)}$$

In the above equations, $x_{i,j}$ is the value of environmental variable *j* in site *i*, ($j = 1, \ldots, n_e$), where n_e is the number of environmental variables. The parameter ψ_s is the maximum expected survival probability in life stage *s*. The functional form in (eqn. 1) maps linear and quadratic combinations of the environmental variables, *u*, to survival, so that survival is always bounded between 0 and ψ_s. The coefficient $\alpha_{1,s,j}$, is the linear trend between the link to survival, of life stage *s*, and the environmental variable *j*, whereas $\alpha_{2,s,j}$ is the quadratic trend. If the quadratic coefficient is negative, $\alpha_{2,s,j} < 0$, survival is maximized at intermediate values of environmental variable *j*, along its gradient, to resemble first principles of the Hutchinsonian niche concept (Holt 2009). Whereas, if $\alpha_{1,s,j} > 0$, and $\alpha_{2,s,j} = 0$, then increases in environmental variable *j* strictly increase survival. Spatial and temporal variation in survival during life stage *s*, not attributable to the environmental variables, $x_{i,j}$ is given by, $\epsilon_{i,s,t}$, a random variable with zero expectation. The intercept, α_0, is set to zero throughout the paper with no loss of generality.

Deleted: The

Deleted: allows for maximum

Deleted: positions along an environmental gradient

Formatted: Font: Italic

Deleted:

Formatted: Font: Italic

Simulated population dynamics - fecundity

As standard in ecological modeling and environmental management (Quinn II and Deriso 1999), we incorporated density-dependence in simulated fecundity. We considered two of the most widely used types of negative density-dependence in the literature. The first was logistic fecundity, where the expected number of juvenile offspring was determined by the logistic model (May 1974). This represented scramble competition, where population sizes above carrying capacity cause declines in total viable offspring. The second was Beverton-Holt (aka. reciprocal yield) density dependence (Shinozaki and Kira 1956, Beverton and Holt 2012). This represented contest competition, decreasing per-capita fecundity, but increasing total fecundity, with respect to population size. Both of these fecundity functions were parameterized with the variables f_{max} and k_i, per capita fecundity at low adult abundance and adult carrying capacity in site i, respectively. We refer to the fecundity function in site i as $f_i(n)$, where n is the number of adults in the given site and time. Note, however, that we did not consider Allee effects and use density-dependence synonymously with strict negative density dependence. For details on the functional forms and parameterization of these standard ecological models, see appendix A.

Simulation algorithm

We considered a population with two life stages, juvenile ($s = 0$) and adult ($s = 1$), and assumed juveniles became adults after one time step or died in that period of time. Therefore, $\varphi(i, 0)$ was the probability of a juvenile in site i transitioning to an adult and, $1 - \varphi(i, 0)$ was the associated mortality probability. The unit of each time step was one iteration of the demographic model, often thought of as one year (Salguero-Gómez et al. 2016). However, for generality, we do not specify the time unit since the species are hypothetical.

Deleted: r

Deleted: -

Deleted: -

Note that a deterministic version of this simulation, without dispersal, is equivalent to simulating a standard matrix population model governed by the density-dependent Lefkovitch matrix (Caswell 2001),

$$\begin{bmatrix} 0 & f_i(n) \\ \varphi(i,0) & \varphi(i,1) \end{bmatrix}, \quad \text{(eqn. 3)}$$

where, $f_i(n)$ is per-capita adult fecundity, and with survival probabilities $\varphi(i,0)$ and $\varphi(i,1)$, determined by eqn 1. Therefore, one can predict the persistence of the population (in an analogous deterministic scenario) by the leading eigenvalue of the linearized system at the extinction equilibrium, e.g. the eigenvalue of,

$$\begin{bmatrix} 0 & f_{max} \\ \varphi(i,0) & \varphi(i,1) \end{bmatrix}, \quad \text{(eqn. 4)}$$

which yields an expected long-term population growth rate, at low population densities, in site, *i*,

$$\lambda_i = \frac{1}{2}\left(\varphi(i,1) + \sqrt{\varphi(i,1)^2 + 4 f_{max}\varphi(i,0)}\right). \quad \text{(eqn. 5)}$$

If $\lambda_i < 1$, a population in site *i* will eventually go extinct without dispersal from other sites or a series of random favorable years; similarly, if $\lambda_i > 1$, the population will persist in the absence of random fluctuations.

The full stochastic simulation, including dispersal, involved four steps. We:

Deleted: ¶

Deleted: *r*

1. Drew the number of surviving juveniles that became adults in the next time step, in each site, from a binomial distribution with the number of trials equal to the number of juveniles in the previous time step and probability of survival, $\varphi(i, 0)$;
2. Drew the number of surviving adults in each site, from a binomial distribution with the number of trials equal to the number of adults in the previous time step and probability of survival, $\varphi(i, 1)$;
3. Drew the number of new juveniles from a Poisson distribution with expectation $f_i(n)n$ [See Appendix A for details about the fecundity function $f_i(n)$]
4. Randomly selected a fraction of new juveniles (offspring in step 3), p_d, to disperse, where each dispersing individual has an equal probability of landing in each site.
5. Updated the total number of adults to equal the surviving adults plus the surviving juveniles (new adults), and updated the total number of juveniles as the new juveniles from reproduction, accounting for offspring entering and exiting the site, due to dispersal.

Simulation scenarios

The main goal of the study was to test how different species varying in survival rates, fecundity, dispersal, responses to environmental variables, stochasticity, and initial abundance affect the accuracy of demographic distribution model (DDM) predictions. We first calculated DDM performance for a baseline scenario, with logistic fecundity, and parameters set to values in Table 1. We then performed a sensitivity analysis, where we simulated the population dynamics and fit distribution models under 1,000 random combinations of parameter values, each under the two different density-dependent fecundity functions, three correlation structures

between fecundity and survival, and three different ways of distributing initial population density in space, each containing approximately 50 different density distributions. This created 1.5 million experiments, each over 4,915 sites and two age-classes (over 145-billion time series).

In both the baseline and the sensitivity analysis, we simulated population dynamics of hypothetical species over 4,195, 10 km^2, sites in Switzerland, affected by elevation, $x_{i,1}$, and forest cover, $x_{i,2}$ (Fig. S1ab), standardized to zero mean and unit standard deviation (Kéry et al. 2017). We let $\epsilon_{i,s,t}$ be independently, identically, normally distributed with mean 0 and standard deviation, σ. Site-specific carrying capacity, k_i, was set to maximum abundance, k_{max}, times the proportion of the site covered by forest. In sites with k_i<5, we set carrying capacity to zero, to represent too little habitat for a persistent, long-term population. We set initial abundances in sites where the species would persist in the absence of stochasticity (sites with $\lambda_i > 1$) to specified proportions of carrying capacity. We randomly selected 5% of sites where forest cover was high enough to yield carrying capacity above 10 individuals, but where survival was too low to maintain a long term viable population (sites with $\lambda_i < 1$) and set their initial adult abundances to 10 individuals, to represent invaded sink populations. There were 200 time-steps in the simulations.

For the sensitivity analysis (see Table 1), maximum adult and juvenile survival, ψ_1 and ψ_0, were assigned random values uniformly drawn between 0.01 and 0.99. This wide range includes slow-growing, long-lived species and fast-growing, short-lived species. The linear effect of forest cover on survival, $\alpha_{1,s,1}$, was randomly drawn from a uniform distribution from 0 to 3. To produce ecologically sensible, yet wide ranges for the effect of elevation on survival, we drew the quadratic elevation effect, $\alpha_{2,s,1}$, from a uniform distribution from -3 to 0, and then also drew a preferred elevation, v, uniformly over the entire elevation range in the data, and chose the

Deleted: initial abundances in sites where the population was expected to persist (sites with $\lambda_i > 1$).

Deleted: 100,000

Deleted: nearly

Deleted: five

linear elevation effect to maximize survival at this preferred elevation, namely, $\alpha_{1,s,1} = -2v\alpha_{2,s,1}$. A quadratic factor of zero represented species that could survive across all observed elevations equally, whereas -3 was for species that could only tolerate a narrow range of elevations.

We considered three scenarios for maximum viable offspring at low population densities, f_{max}, (1) positively correlated with survival, (2) negatively correlated with survival, and (3) fixed across sites. For the fixed case, in each of the 1,000 parameter sets, f_{max} was randomly drawn from a uniform distribution ranging between the lowest possible number such that the species would be expected to persist in at least 5% of sites (i.e. $\lambda_i > 1$), and the largest possible number for which carrying capacity was guaranteed to be a stable equilibrium at the most favorable site, given ψ_1, ψ_0, and the effects of environmental variables (which were drawn first). The last constraint simply eliminated the possibility of chaotic and unstable, periodic dynamics, and was determined through standard linear stability analysis (Strogatz 1994) techniques (see Appendix B). In the correlated case, f_{max}, was different at each site according to the environmental variables at that site. This was achieved by drawing a number between the maximum and minimum values for f_{max} described above, for each site, using a beta distribution, with beta distribution parameters as a function of environmental variables. This made f_{max} more likely to be high in sites with high survival, and low in sites with low survival (see the section "environmentally driven fecundity scenarios" in Appendix B for details). The negatively correlated scenario was achieved similarly (Appendix B).

The standard deviations of environmental stochasticity and the dispersal proportion were uniformly randomly generated on (0, 0.5) and (0, 0.05), respectively, to represent wide ranges for the types of species an ecologist would consider fitting a deterministic demographic model

Deleted: Maximum

Deleted: *r*

without dispersal. Maximum carrying capacity across all sites, k_{max}, was also uniformly distributed.

We considered three different scenarios for how initial population sizes were distributed through space: (1) fixed initial abundances at all sites where the species was expected to persist were varied factorially with each parameter combination, using 51 values between 5% and 135% of local carrying capacity; (2) initial abundance set at 25, 50 and 75% of the carrying capacity in a fixed proportion of sites and carrying capacity in the other sites (varied across all possible proportions); and (3) uniformly distributed abundance, with 46 mean values [from 52.5%, corresponding to a lower bound of 5%, to 100% of carrying capacity].

Deleted: ,

Table 1. A list of parameters used in the species range dynamics baseline simulations (third column), which are perturbed in the sensitivity analysis (fourth column) to test the generality of results for different types of hypothetical species.

Parameter	Description	Baseline Value	Range tested
ψ_0	Maximum juvenile survival	0.5	0.01 – 0.99
ψ_1	Maximum adult survival	0.5	0.01 – 0.99
$\alpha_{1,s,1}$	Linear forest cover effect of stage *s* survival	0.5	0 – 3
$\alpha_{1,s,2}$	Linear elevation effect on stage *s* survival	0.5	-5 – 23*
$\alpha_{2,s,2}$	Quadratic elevation effect on stage *s* survival	-0.5	-3 – 0
p_d	Dispersal proportion per time step	0.01	0 – 0.05
σ	Standard deviation of environmental stochastic effect on survival	0.25	0 – 0.5
f_{max}	Fecundity at zero density, i.e. maximum fecundity [surviving offspring / adult]	4	0.1 – 250
k_{max}	Maximum carrying capacity (at a site with 100% forest cover) [number of adults]	1,000	500 – 1,000

*chosen to maximize survival at a uniform randomly chosen preferred elevation, over all possible elevations in the data, given a random quadratic elevation effect, drawn from the range in the row below.

Deleted: *r*

Simulated field sampling

We assumed that simulated population dynamics represented the true population, and sampled from this population by simulating a demographer using common field-sampling methods (Zurell et al. 2010). Sampling occurred for ten time-steps from the start of the simulation with dispersal turned off, to simplify the analysis and mimic a situation where the researcher can account for the origin of individuals in the site. The virtual ecologist randomly chose n_s sites where the species was present at the beginning of the simulation and then counted the number of surviving juveniles (new adults), surviving adults, and new juveniles, at each sampled site, over ten time-steps. While this is a standard approach (Lavine et al. 2002), an alternative, but more laborious and computationally expensive, approach, would model individual organisms, and track a sampled subset of these individuals. Tracking individuals is advantageous if one wants to quantify individual variability in demographic processes, but this was not a focus of our study. Also, considering we were analyzing nearly five-billion time series, computational efficiency was required to make sure results were general across species. We set n_s = 200 sites, representing a highly optimistic, but realistic sample size. For example, previously, DDMs have used 138 sites (Merow et al. 2014). We selected a high value because the purpose of the study was to identify species for which DDMs could generate useful predictions given high-quality data, but we also tested DDM accuracy for scenarios with 50, 30, 25 and 20 sampled sites. We set the length of demographic surveys to two years, but we also tested survey lengths of three, five, 10 and 20 years.

Distribution models

The demographic distribution model (DDM) assumed that the population dynamics in site *i* were governed by a two-stage matrix population model. While a variety of density-independent demographic models have been used to build DDMs in the literature, including integral projection models (Merow et al. 2014, 2017) and matrix population models (Buckley et al. 2010), we chose a matrix approach for both the simulation and fitted DDM because it is the simplest and most computationally efficient model that maintains the essential demographic features of structured population dynamics. The fitted model was,

$$\boldsymbol{N}_{i,t+1} = \boldsymbol{A_i}\boldsymbol{N}_{i,t};$$

(eqn. 6)

$$\boldsymbol{A}_i = \begin{bmatrix} 0 & f(x_{i,1}, x_{i,2}) \\ \phi_0(x_{i,1}, x_{i,2}) & \phi_1(x_{i,1}, x_{i,2}) \end{bmatrix},$$

where $\boldsymbol{A_i}$ is a transition matrix for site *i*, and $f(x_{i,1}, x_{i,2})$, $\phi_0(x_{i,1}, x_{i,2})$, and $\phi_1(x_{i,1}, x_{i,2})$ are the estimated fecundity, juvenile and adult survival in site *i*, respectively. Note these are functions of the two environmental variables in site *i*, $x_{i,1}$, and $x_{i,2}$. The estimation of $f(x_{i,1}, x_{i,2})$, $\phi_0(x_{i,1}, x_{i,2})$, and $\phi_1(x_{i,1}, x_{i,2})$ was performed using a statistical model. From the simulated field sampling of demographic data, we computed the observed per-capita fecundity, number of surviving adults and juveniles over the sampling period. These quantities were uniquely determined because we turned off dispersal during the period of demographic sampling. This procedure created a vector of observed juveniles survived, adults survived, and fecundity, each with the number of entries equal to the number of sites. The functions ϕ_0, ϕ_1, and f, were then estimated using generalized additive models, function 'gam' in R (Wood 2017). The generalized

Deleted: was

Deleted: survival proportion,

Deleted: juvenile survival proportion

Deleted: by tracking the number of new juveniles, new adults and surviving adults

Deleted: survival

Deleted: survival

Deleted: data

Deleted: (

Deleted:)

Deleted: with Gaussian errors

Formatted: Highlight

additive models assumed binomially distributed counts of surviving adults and juveniles and Poisson distributed total offspring with a rate parameter equal to an estimated parameter, based on environmental predictors, times the number of adults at the site. The estimated parameter was therefore expected per-capita fecundity. We restricted total offspring predictions to the range of observed values to avoid issues extrapolating beyond the data, as is common in distribution modelling (Stohlgren et al. 2011, Owens et al. 2013).

The DDMs predicted a unique matrix, $\boldsymbol{A}_i$, for every site based on the environmental variables at that site. We used the predicted $\boldsymbol{A}_i$ to calculate the long-term population growth rate, λ_i, by computing $\boldsymbol{A}_i$'s leading eigenvalue. Following standard practice for DDMs (Merow et al. 2014, 2017), we interpreted λ_i as a measure of persistence, where $\lambda_i<1$ predicted eventual extinction and $\lambda_i>1$ predicted long-term persistence at a given site. We then compared the predicted λ_i values to the presence of the species, 200-time steps after demographic sampling, to determine whether DDM predictions of persistence were correlated with long-term persistence at a site. Note that, throughout the paper, we refer to λ_i as long-term population growth rate from the fitted DDM, whereas λ_i is the expected population growth rate from a deterministic version of the true model used to simulate the data.

We also compared how accurately the DDMs predicted occupancy in comparison to correlative species distribution models (SDMs). SDMs were generalized additive models, predicting the probability of species' presence given presence/absence data and environmental variable values, at the n_s sampled sites, at the end of demographic sampling. In cases where the generalized additive models did not converge (less than 0.1 percent of scenarios), for both the DDM and SDM, we used generalized linear models with similarly distributed error. To calculate

Deleted: normally
Deleted: fecundity
Deleted: offspring
Deleted: proportional to
Deleted: „ and logit transformations for ϕ_0, and ϕ_1
Deleted: (Wood 2017)
Deleted: . Survival data was coerced between 0.0001 and 0.9999 instead of zero and one to prevent division by zero or taking the log of zero in the logit transformation
Formatted: Highlight
Formatted: Highlight
Deleted: linear
Deleted:)
Deleted: The
Deleted: a
Deleted: model withSDMs assumed a binomially distributed response andwith a logit link function

the prediction accuracy of the SDM we considered sites to be predicted present when the modeled probability of presence was greater than 0.5.

Results

For the baseline scenario, with all parameters set to intermediate values, and initial population sizes in each site set to 10% of local carrying capacity, the DDM performed well. For 95.1% of the 3,518 sites where the population went extinct (light grey in Fig. 2a), the DDM predicted $\lambda_i<1$ (Fig. 2b). For the 1,397 sites where the species was present at the end of the simulation (green in Fig. 2a), the DDM predicted $\lambda_i>1$ in 99.7% of these sites. The lack of red pixels in Fig. 2c denotes the 0.3% of present sites where the DDM incorrectly predicted $\lambda_i<1$.

In the same baseline scenario, but with initial populations at carrying capacity during the start of demographic sampling, the DDM over-predicted extinction (Fig. 2 d-f). For the 1,415 sites where the species was present at the end of the simulation, the DDM predicted $\lambda_i>1$ at only 39.0% of the sites. On the other hand, out of the 3,500 sites where the population went extinct (light grey in Fig. 2d), the DDM correctly predicted $\lambda_i<1$, 97.3% of the time (Fig. 2d). This means the DDM predicted the status of absent sites well regardless of population density at demographically sampled sites.

Deleted: 6

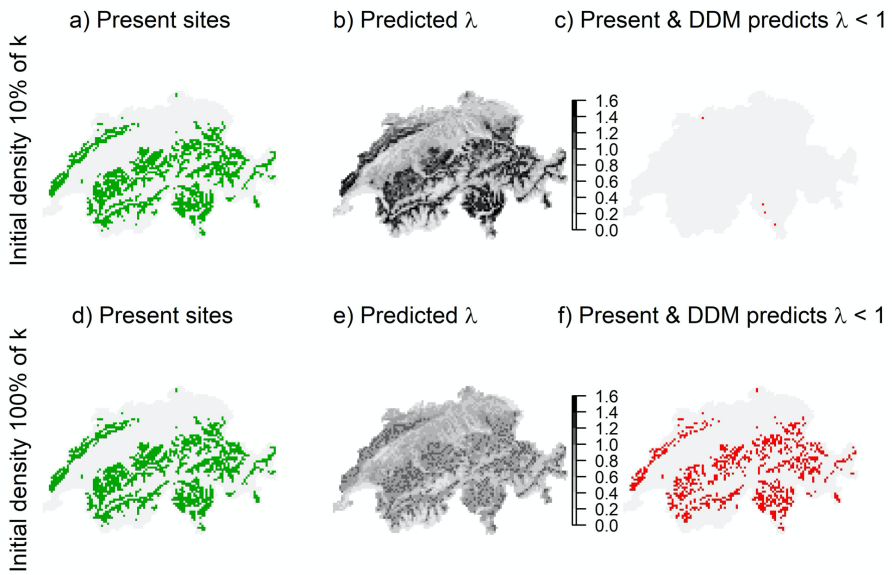


Fig. 2. Maps of occupancy and predicted population growth rate, λ, from the Demographic Distribution Model (DDM) of a virtual species in Switzerland, given demographic data sampled from sites with low (top row) and high (bottom row) population densities, showing an over prediction of extinction when demographic data come from locations at carrying capacity. (a, d) Map of occupancy at the end of the simulation, (b, e) predicted population growth rate λ from the DDM, and (c, f) sites where present populations at the end of the simulation are incorrectly predicted by the DDM to be absent. The initial population sizes at sampled sites were at 10% of carrying capacity in row 1 (a-c) and at 100% of carrying capacity in row 2 (d-f), for the baseline parameterization. When demographic samples were conducted at sites where the populations were at 10% of carrying capacity, the DDM correctly predicted all present sites as present, whereas the DDM only correctly predicted 56.1% of present sites when demographic sampling occurred at carrying capacity.

To summarize, in the baseline scenario, DDMs predicted present sites accurately if initial density during sampling was close to zero, and inaccurately for initial density at carrying capacity. However, for what initial density in sampled sites, between 10% and 100% of carrying capacity, do predictions cease to be accurate at present sites? To identify the critical population density to achieve a specified target percentage of correctly predicted present sites, we ran the above simulation with initial densities of 5% to 135% of carrying capacity (Fig. 3) and computed the proportion of present sites where the DDM predicted $\lambda_i>1$ for each initial density (black dots in Fig. 3). The black dots in Fig. 3 formed a clear monotonic decreasing pattern, and we fit a smooth curve to these data (curve in Fig. 3, see Appendix C for details on curve fitting methods). We then computed the critical population abundance as the intersection of this curve with the specified target prediction accuracy (Fig. 3). A critical initial population size of 94.2% of carrying capacity was required to achieve a prediction accuracy of 80% in present sites (circle in Fig. 3). The critical population abundance needed to correctly predict 90% of present sites was 90.1% of carrying capacity (square in Fig. 3). Unlike true presence predictions, correctly predicting absent sites was not strongly related to population density, which we discuss further in the results of the sensitivity analysis.

Deleted: 2
Deleted: 2
Deleted: 2
Deleted: 2
Deleted: 89.1
Deleted: 2
Deleted: 74.7
Deleted: 2

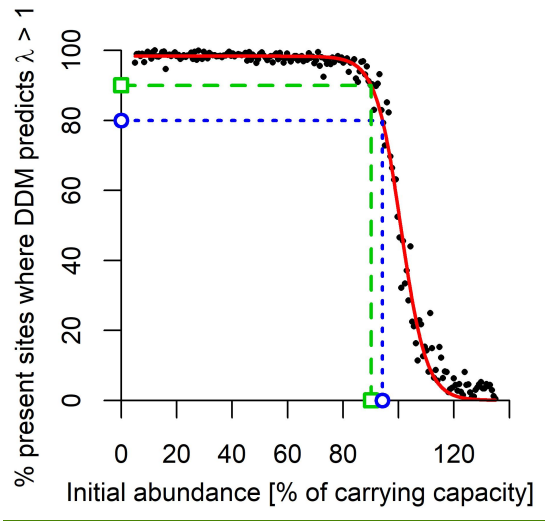


Fig. 3. The probability that the DDM correctly predicted species' presence at present sites declined as population size, relative to carrying capacity, increased at sampled sites. Each black dot is the proportion of occupied sites at the end of a single simulation where the DDM predicted $\lambda_i>1$, given an initial population size, during demographic sampling, specified by the x-axis. The red line is a fitted smooth monotonic curve to these data. The green square and blue circle on the x-axis are the critical initial population sizes (90.1%, and 94.2% of carrying capacity, respectively) required, during demographic sampling, to achieve a 90% and 80% chance of correctly predicting occupied sites at the end of the simulation. All parameters were set to their baseline, and the fecundity function was logistic. Absent sites were always predicted absent by the DDM with 96-100% accuracy regardless of initial population abundance in sampled sites and hence are not displayed here.

Deleted: 74.753

Deleted: 89.13

To demonstrate that Fig 3 was not an artefact of the baseline parameterization chosen, we then proceeded to calculate the critical population abundance, as in Fig. 3, for every parameter combination, fecundity function, and correlation scenario in the sensitivity analysis. We set the true presence accuracy threshold equal to 80% accuracy, a round number close to the 79% prediction accuracy reported from past empirical DDMs (Merow et al. 2014). A histogram of critical population abundances across all parameter combinations tested, under both fecundity functions and all three correlation scenarios, shows that high critical initial population sizes (80-95% of carrying capacity) at demographically sampled sites were the most common (Fig. 4). This means that DDMs predicted presence accurately even when abundances at sampled sites were intermediate rather than small. However, in the fixed fecundity scenario, when maximum fecundity was not spatially correlated with survival, the DDMs did not accurately predict occupancy at present sites for a few parameter combinations, even when sites had small population sizes during sampling (small leftmost bar in the middle column of Fig. 4).

Deleted: ure
Deleted:
Deleted: 2
Deleted: each
Deleted: (see Fig. 2, and Fig. S2 for the first 6 curves in the sensitivity analysis)
Deleted: 75
Deleted: , for a few parameter combinations
Deleted: s

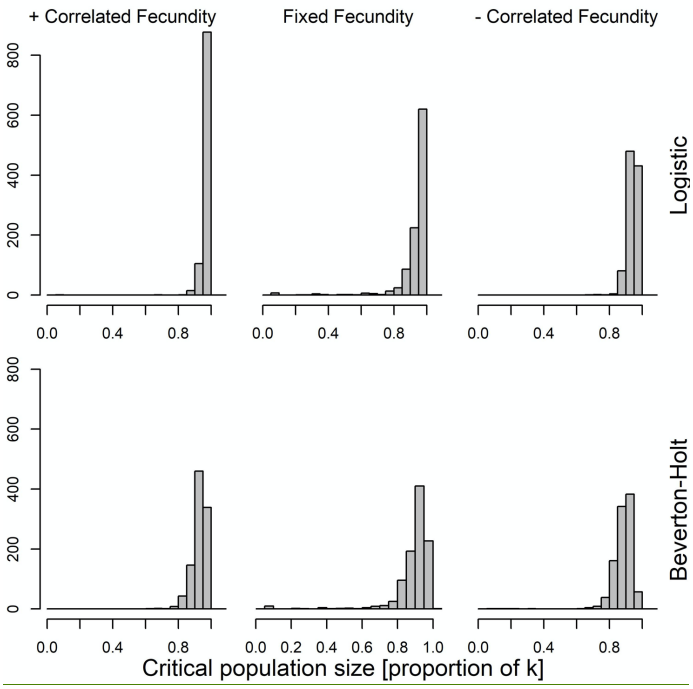
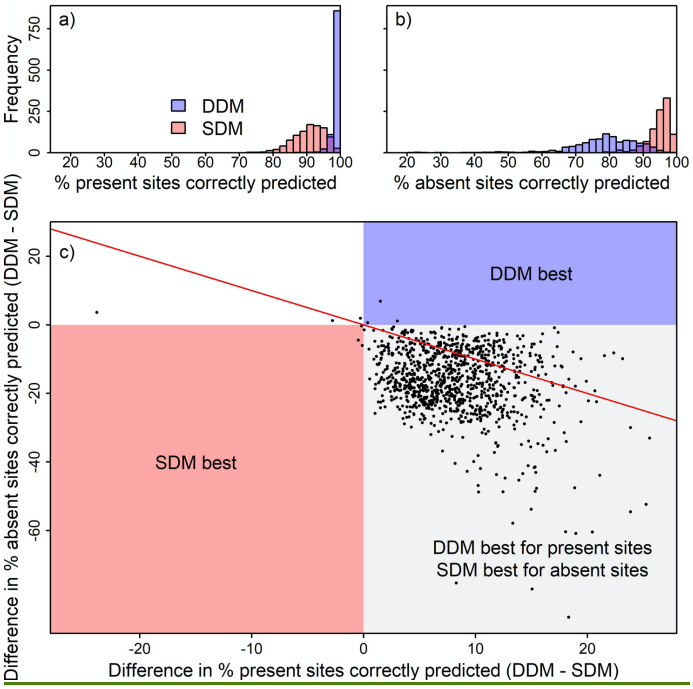


Fig. 4. Histogram of critical population size at the start of demographic sampling required for the DDM to achieve 80% prediction accuracy at present sites, across 1,000 randomly drawn parameter combinations. Critical population density [initial abundance as a proportion of carrying capacity] was typically between 0.75 - 0.95. The top and bottom rows are for Logistic and Beverton-Holt negative density-dependent fecundity functions. Maximum fecundity, at low population densities, is positively and negatively correlated with survival in the left and right columns. In the middle column, maximum fecundity is fixed across the landscape. As long as both survival and fecundity were correlated with environmental variables, DDMs achieved an 80% prediction accuracy in nearly all simulations. Note that critical population size means the population must be at or lower than this population size, to achieve the accuracy threshold.

Deleted: 3

In the case where populations were at 50% of carrying capacity at the start of demographic sampling DDMs had higher prediction accuracy at present sites than correlative species distribution models (SDMs) in both the correlated (Fig. 5a) and negatively correlated (Fig. S2a) fecundity scenarios. However, SDMs predicted locations where the population was absent more accurately (Fig. 5b and S2b). The trade-off in improved accuracy at present and absent sites for DDMs and SDMs, respectively, occurred in over 98 percent of the simulations (points in the lower-right, grey region in Fig 5c and S2c). SDMs improved total accuracy over DDMs more frequently if survival was correlated with fecundity (more points below the 1-1 line in Fig 5c). However, the DDM improved overall accuracy if survival and fecundity were negatively correlated (more points above the 1-1 line in Fig S2c). In the negatively correlated fecundity scenario, the accuracy of both methods declined compared to the correlated scenario (see more left bars in the histograms in Figs S2ab than 5ab). However, the relative improvement of DDMs over SDMs was due to major decreases in SDM accuracy at present sites under negatively correlated fecundity (see the leftward shift of pink bars in Fig S2a compared to Fig 5a).

Deleted: DDM prediction accuracy at present sites was correlated with the long-term population growth rate of the species at low densities, λ_i, in sites where the species occurred. In Fig. 3cd, critical population abundance was usually 65-95% of carrying capacity, as long as the median of λ_i in present sites was greater than 1.15 (Fig. 3cd). For slow-growing species, median λ_i less than 1.1, DDMs were likely to over-predict extinction, even if demographic data was sampled at sites well below carrying capacity (left-most points in Fig. 3cd). No other parameters in the simulation strongly correlated with the critical population abundance at sampled sites (Fig. S3). These results were robust to the length of demographic sampling (see Fig. S4 – S9) but were even more pronounced when reducing the sampling period from 10 to six timesteps (Fig. S4-S6).¶
SDMs accurately predicted occupancy, and there was no strong trend between SDM prediction accuracy and population growth rate (first column in Fig. 4). Instead, SDM prediction accuracy, at present sites, was influenced by the number of present sites at the end of the simulation (shading in Fig. 4a). The two outliers (light, lower-left points in Fig. 4a), were the only two simulation runs where the species persisted in less than 5% of sites.¶
Alternatively, DDM presence prediction accuracy was much more strongly correlated with population growth rate. DDMs over-predicted extinction for slow-growing populations, even when the initial density was at 5% of carrying capacity in sampled sites (Fig. 4b). For fast growing-populations, the DDM more frequently over-predicted presence at absent sites (Fig. 4d). For the SDM, the opposite was true, there was a slight tendency to over-predict absence at present sites for faster-growing species.



574

575 **Fig. 5.** DDM and SDM prediction accuracy at (a) present and (b) absent sites, and (c) difference
576 in prediction accuracy between SDMs and DDMs at present sites (horizontal axis) and absent
577 sites (vertical axis), given initial densities at 50% of carrying capacity. Positive values indicate
578 that DDMs have higher prediction accuracy. In (c) each point corresponds to a single randomly
579 sampled parameter set. Almost all points fall in the lower right quadrant, corresponding to
580 parameter sets where DDMs more accurately predicted present sites, and SDMs more accurately
581 predicted absent sites. Points above the red one-to-one line correspond to parameter sets where a
582 DDM's improved prediction accuracy at present sites is higher than the DDM's decreased
583 prediction accuracy at absent sites. Fecundity was logistic, and positively correlated with
584 survival.

All of the results discussed thus far assumed that populations were at a specified density, relative to carrying capacity, at all sites used to build the DDM. If only a portion of sites started at a specified density, perturbed below carrying capacity, while others started at carrying capacity, predictions worsened (Fig S3, S4). For example, if fecundity was logistic, 25 – 61 percent of sites supplying demographic data needed to be perturbed from carrying capacity in order to achieve 80 percent prediction accuracy in 90 percent of the simulations (Fig S3) across the three fecundity correlation structures and three perturbation magnitudes tested. The results were qualitatively similar for Beverton-Holt fecundity, but with a higher proportion of sites that needed to be perturbed (34-75% of sites) to achieve the same accuracy (Compare Fig S3a-c to Figs. S4a-c). In the case where initial population size was uniformly distributed at each site, 80 percent prediction accuracy was achieved if mean initial densities exceeded 71% of carrying capacity, across all scenarios (Fig S4).

In general, given 200 sampled sites, across all scenarios, if mean initial abundance, averaged across sites, was 70% of carrying capacity or less, a density-independent DDM predicted present sites with at least 80% accuracy, in at least 90% of the simulations (Fig S3-S6, and 6a). However, demographic data from fewer than 200 sites, meant more sites had to be perturbed below carrying capacity during demographic sampling to achieve a desired DDM prediction accuracy. For example, for logistic fecundity, positively correlated with survival, and populations at 50% of carrying capacity in perturbed sites, the DDM required 25, 61, and 87% of sites to be perturbed below carrying capacity, given 200, 50, and 30 sampled sites, respectively (Fig. 6a-c). Even if all populations were at 50% of carrying capacity during the start of demographic sampling, it was impossible to guarantee 80% accuracy in 90% of the simulations, if there were 25 sampled sites or fewer (Fig. 6de). Additionally, DDMs built with long-term

Deleted: 6

Deleted: if perturbed population density was 25 percent of carrying capacity, then 16

Deleted: 6a

Deleted: For intermediate and small perturbed population sizes of 50% and 75% of carrying capacity, 21 and 38 percent of sites needed to be perturbed, respectively (Fig 6bc). …

Deleted: all combinations of density dependent fecundity functions, and correlation scenarios (Figs. S3 and S4). However, …

Deleted: led to

Deleted: below carrying capacity, 20, 35, and 80

Deleted: percent of sites

Deleted: and in the 25%, 50%, and 75% of carrying capacity scenarios, respectively

Deleted: 4

Deleted: 6 and

Deleted: 3

Deleted: across Beverton-Holt and fixed correlation scenarios, only for

Deleted: close

Deleted: to

Deleted: 5

Deleted: 0

Deleted: for all distributions of initial abundance tested,

Deleted: was at

Deleted: 50

Deleted: , averaged across all demographically sampled sites, …

Deleted: 6 and

Deleted: 5

Deleted: This held true as long as there were several sites with available demographic data (>20, Fig S6), and the sampling interval was short (2-5 time steps, see Fig S7).

monitoring data (e.g. 10 time-steps or greater) provided less accurate predictions (right column of Fig S6) despite increased sample sizes, because sampled populations frequently saturated at carrying capacity during data collection.

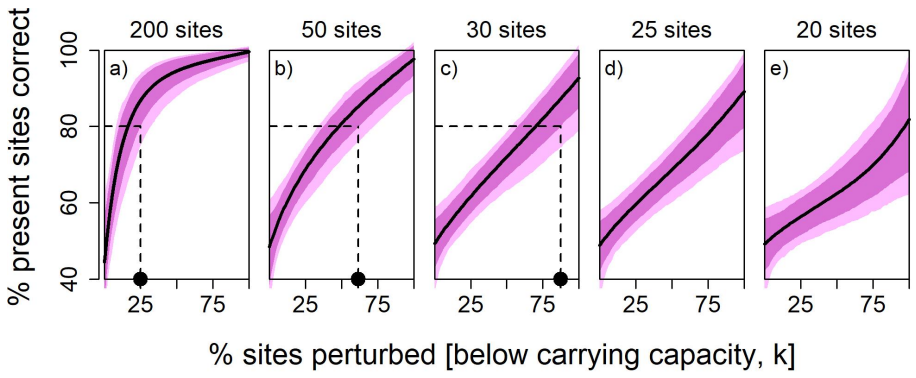


Fig. 6. Percent of present sites correctly predicted by the DDM as a function of the percentage of sites perturbed 50% below carrying capacity given data from (a) 200, (b) 50, (c) 30, (d) 25, and (e) 20, sites respectively. Dark and light shaded regions are 80 and 95% confidence intervals, respectively. Black circles are the critical percentage of sites that must be perturbed below carrying capacity during demographic sampling to achieve 80% prediction accuracy or higher at present sites in 90% of the simulations. As the number of sampled sites decreases, a higher percentage of sites need to be below carrying capacity to achieve the desired accuracy. If there were 25 sampled sites, or fewer, it was not possible to achieve 80% prediction accuracy, in 90% of the simulations (d, e). Fecundity was logistic, and maximum fecundity was positively correlated with survival.

Deleted: when perturbed populations were allowed to grow over longer periods

Deleted: ¶

Deleted: ;

Deleted: ;

Deleted: .

Deleted: „ 38,16(b), 21

Deleted: (for fixed and scenarios see Fig S, and for Beverton-Holt fecundity see Fig S).

Formatted: Font: Not Bold

Discussion

Using 1.5 million simulations of hypothetical species' range-dynamics, we evaluated when deterministic, density-independent, demographic distribution models (DDMs) accurately predicted species' distributions. While DDM predictions were biased towards species absence in simulations where data used to build the DDM came from sites with populations near carrying capacity, our comprehensive simulations support the following generality: if mean initial population size is less than 50% of carrying capacity, averaged across >25 sites, where demographic data is available, a density-independent DDM will predict present sites accurately. When this condition was satisfied, DDMs outperformed correlative species distribution models (SDMs) at accurately predicting present sites, but, not absent sites. Our results suggest that density-independent DDMs may be useful for predicting species ranges for some taxa and applications (Briscoe et al. 2019), despite obvious limitations (Ellner et al. 2016). Often species' have multiple locations available for sampling with populations well below carrying capacity (Haak 2000, Williams et al. 2011), and there are 11 species of plants and animals with matrix population models built from data at >25 sites, already available in the global, open-access databases COMPADRE and COMADRE (Salguero-Gómez et al. 2015, 2016).

One particularly important application of density-independent DDMs is invasive species risk mapping (Merow et al. 2017), an application where traditional SDM approaches have been criticized (Liu et al. 2020). Our work confirms that this is potentially an appropriate use of DDMs because invasive species, at the start of an invasion, are typically well below carrying capacity (Ramula et al. 2008, Davis 2009, Burns et al. 2013). Additionally, because successful invaders often invade multiple locations across wide geographic ranges, they are also species for which demographic data replicated across geographically distant and climatically dissimilar sites

Deleted: of thousands

Deleted: , or if populations grew slowly.

Deleted: rule of thumb. I

Deleted: 0

Deleted: even

Deleted: a

Deleted: tool

Deleted: a broad range of

Deleted: their

Deleted: 16

Deleted: 0

Field Code Changed

Deleted: , and also usually grow quickly

can be collected (Marshall 2016). Our simulations suggest more than 25 such sites are required for accurate predictions. While this will be feasible for many invasive species (Merow et al. 2017), DDMs will not be useful for mapping invasion risk for species already abundant throughout their entire invaded ranges or those whose populations start small but quickly saturate to carrying capacity during demographic data collection.

DDMs may also be appropriate for predicting threatened species ranges, because threatened populations are often at low densities (IUCN 2012). Since DDMs tie these predictions to mechanisms of decline and growth, they may provide insight into which management actions will maximize species persistence (Briscoe et al. 2019). However, our results identify a key limitation of using DDMs for threatened species management. While DDMs accurately predicted presence at sites where the species persisted, correlative species distribution models SDMs outperformed DDMs at sites where the species was absent in over 95 percent of the simulations. Therefore, our results suggest that DDMs may be more appropriate for applications where identifying present sites correctly is more important than identifying absent sites. For example, in invasion risking mapping, predicting a local invasion at a site where an invasion fails to occur is a more tolerable error than missing the location of a future invasion. In this case, present site predictions are more important and a DDM will be appropriate, if demographic data from low density sites are available. In contrast, a manager looking for a site to release a threatened species, to establish a new population (often called a species "translocation"), would not want to select a site where the species will go extinct. Therefore, when identifying translocation sites, predicting absences accurately is very important, and DDMs may be inappropriate. Our DDMs possibly overpredicted presence in sites where the species went extinct because they were deterministic and therefore did not predict extinctions caused by demographic stochasticity.

Deleted: -

Deleted: But

Deleted: invasion risk maps are likely

Deleted: less

Deleted: the many

Deleted: are

Deleted: in

Deleted: range expansions of quickly recovering

Deleted: after their threats have been abated. This is

Deleted: species

Deleted: and reducing threats via conservation actions, such as habitat restoration (Garnett et al. 2019), removal of invasive predators or competitors (Legge et al. 2011, Kanowski et al. 2018, Prior et al. 2018), and harvest bans (Belcher and Jr. 2002, Fukuda et al. 2011), have led to high population growth rates during the recovery of several threatened populations

Deleted: seed

Formatted: Highlight

Applications for which correct predictions in unsuitable habitat are more important than predictions in suitable habitat, therefore, require a stochastic DDM, or a coupled niche-population model, to account for demographic stochasticity explicitly (Keith et al. 2008, Fordham et al. 2013), or an SDM, which can account for extinctions from demographic stochasticity implicitly.

Perhaps the biggest obstacle for wide-scale use of density-independent DDMs is that, for many species, demographic data are sampled primarily at sites where the species is already abundant (Quintana-Ascencio et al. 2018, Fournier et al. 2019). Field ecologists do not typically risk designing randomized, labor-intensive, demographic studies at sites where populations are so small that the ecologists may not even find individuals to sample. However, there is a small subset of the demography literature focused on populations at the edge of species ranges, where population density is often small (Sexton et al. 2009, Eckhart et al. 2011, Pironon et al. 2017). Our results show that demographic data from such low-density sites are highly valuable for building accurate density-independent DDMs.

There are three major caveats behind our approach to identifying guidelines for when to use density-independent DDMs. First, we simulated range dynamics using a simple model with several underlying assumptions, such as common pool dispersal, random/fixed initial population densities across sites, density dependence in fecundity, and demographic rates influenced by life stages rather than continuous traits (Easterling et al. 2000). Second, we assumed no systematic environmental change (e.g. climate change or deforestation). Lastly, the DDM used a model that closely matched the stochastic model employed to simulate the data. Some of these choices may have affected the relative performance of DDMs *vs.* SDMs. For example, DDM performance may have improved, relative to SDMs, if we included systematic, environmental change,

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Deleted: This is the exact opposite case .F at a site a error the location of a futureThis provides another reason why DDMs are a promising tool for invasion risk mapping.

allowing estimated vital rates to change through time with environmental drivers (Evans et al. 2016). And the relative accuracy of DDMs might have declined if the DDM model did not closely match the simulation model. However, the main result of our paper, that density-independent DDMs will only be useful when demographic data are collected from areas where populations are below carrying capacity is likely robust to all of these caveats.

For modelers who wish to predict species ranges using process-explicit models for species with data collected from populations near carrying capacity, one option is to model density dependence explicitly. Simulation studies have found that models including carrying-capacity (Pagel and Schurr 2012, Schurr et al. 2012) can outperform density-independent methods when predicting species occurrence (Zurell et al. 2016). And there are a few examples of empirically-driven density-dependent demographic models (Vanderwel et al. 2013, García-Callejas et al. 2016, Pagel et al. 2020). These models, which incorporate density dependence explicitly are likely required to predict transient dynamics and project species abundances (rather than just presence absence). Unfortunately, fitting density-dependent models requires, not only data at multiple sites, across multiple environmental conditions, and tracking multiple life stages, but also requires replication across populations at different densities. This may be impractical to obtain in many scenarios. Here, we have demonstrated that ignoring density dependence when predicting species ranges from demographic models is a practical first step and likely appropriate in several situations.

Deleted: two
Deleted: s
Deleted: , or when species grow quickly, are
Deleted: project
Deleted: slow-growing species, or
Deleted: explicitly
Deleted: projecting

797 **References**

798 Ackerly, D. D. et al. 2010. The geography of climate change: implications for conservation
799 biogeography. - Divers. Distrib. 16: 476–487.
800 Barbraud, C. et al. 2011. Contrasted demographic responses facing future climate change in
801 Southern Ocean seabirds. - J. Anim. Ecol. 80: 89–100.
802 Beverton, R. J. H. and Holt, S. J. 2012. On the Dynamics of Exploited Fish Populations.
803 Briscoe, N. J. et al. 2019. Forecasting species range dynamics with process-explicit models:
804 matching methods to applications. - Ecol. Lett. 22: 1940–1956.
805 Buckley, L. B. et al. 2010. Can mechanism inform species' distribution models? - Ecol. Lett. 13:
806 1041–1054.
807 Burns, J. H. et al. 2013. Greater sexual reproduction contributes to differences in demography of
808 invasive plants and their noninvasive relatives. - Ecology 94: 995–1004.
809 Cabral, J. S. et al. 2017. Mechanistic simulation models in macroecology and biogeography:
810 state-of-art and prospects. - Ecography (Cop.). 40: 267–280.
811 Caswell, H. 2001. Matrix population models : construction, analysis, and interpretation. - Sinauer
812 Associates.
813 Csergő, A. M. et al. 2017. Less favourable climates constrain demographic strategies in plants (J
814 Gurevitch, Ed.). - Ecol. Lett. 20: 969–980.
815 Cushing, J. M. et al. 2002. Chaos in ecology : experimental nonlinear dynamics. - Academic
816 Pres.
817 Dahlgren, J. P. et al. 2014. Local environment and density-dependent feedbacks determine
818 population growth in a forest herb. - Oecologia 176: 1023–1032.
819 Davis, M. A. 2009. Invasion Biology. - Oxford University Press.

820 Diez, J. M. et al. 2014. Probabilistic and spatially variable niches inferred from demography (E
821 Jongejans, Ed.). - J. Ecol. 102: 544–554.
822 Easterling, M. R. et al. 2000. Size-Specific Sensitivity: Applying a New Sructured Population
823 Model. - Ecology 81: 694–708.
824 Eckhart, V. M. et al. 2011. The geography of demography: long-term demographic studies and
825 species distribution models reveal a species border limited by adaptation. - Am. Nat. 178
826 Suppl 1: S26-43.
827 Ehrlén, J. et al. 2016. Advancing environmentally explicit structured population models of plants
828 (A Griffith, Ed.). - J. Ecol. 104: 292–305.
829 Elith, J. and Leathwick, J. R. 2009. Species Distribution Models: Ecological Explanation and
830 Prediction Across Space and Time. - Annu. Rev. Ecol. Evol. Syst. 40: 677–697.
831 Ellner, S. P. et al. 2016. Density Dependence. - In: Data-driven Modelling of Structured
832 Populations. Springer, Cham, pp. 111–138.
833 Evans, M. E. K. et al. 2016. Towards Process-based Range Modeling of Many Species. - Trends
834 Ecol. Evol. 31: 860–871.
835 Fordham, D. A. et al. 2013. Tools for integrating range change, extinction risk and climate
836 change information into conservation management. - Ecography (Cop.). 36: 956–964.
837 Fordham, D. A. et al. 2018. How complex should models be? Comparing correlative and
838 mechanistic range dynamics models. - Glob. Chang. Biol. 24: 1357–1370.
839 Fournier, A. M. V. et al. 2019. Site-selection bias and apparent population declines in long-term
840 studies. - Conserv. Biol. 33: 1370–1379.
841 García-Callejas, D. et al. 2016. Projecting the distribution and abundance of Mediterranean tree
842 species under climate change: a demographic approach. - J. Plant Ecol. 10: rtw081.

843 Guisan, A. et al. 2013. Predicting species distributions for conservation decisions (H Arita, Ed.).
844 - Ecol. Lett. 16: 1424–1435.
845 Haak, C. 2000. The concept of equilibrium in population ecology.
846 Hijmans, R. J. et al. 2017. dismo: Species distribution modeling.: 1–68.
847 Holt, R. D. 2009. Bringing the Hutchinsonian niche into the 21st century: ecological and
848 evolutionary perspectives. - Proc. Natl. Acad. Sci. U. S. A. 106 Suppl 2: 19659–65.
849 IUCN 2012. IUCN Red List Categories and Criteria. - IUCN.
850 Keith, D. A. et al. 2008. Predicting extinction risks under climate change: coupling stochastic
851 population models with dynamic bioclimatic habitat models. - Biol. Lett. 4: 560–563.
852 Kéry, M. et al. 2017. AHMbook: Functions and Data for the Book "Applied Hierarchical
853 Modeling in Ecology."
854 Lavine, M. et al. 2002. Statistical modeling of seedling mortality. - J. Agric. Biol. Environ. Stat.
855 7: 21–41.
856 Liu, C. et al. 2020. Species distribution models have limited spatial transferability for invasive
857 species (C Bertelsmeier, Ed.). - Ecol. Lett.: ele.13577.
858 Marshall, G. 2016. Individual performance and population dynamics of Plantago lanceolata in
859 Southeast Queensland.
860 May, R. M. 1974. Biological Populations with Nonoverlapping Generations: Stable Points,
861 Stable Cycles, and Chaos. - Science (80-.). 186: 645–647.
862 Merow, C. et al. 2014. On using integral projection models to generate demographically driven
863 predictions of species' distributions: development and validation using sparse data. -
864 Ecography (Cop.). 37: 1167–1183.
865 Merow, C. et al. 2017. Climate change both facilitates and inhibits invasive plant ranges in New

866 England. - Proc. Natl. Acad. Sci. U. S. A.: 201609633.
867 Meyer, C. et al. 2015. Global priorities for an effective information basis of biodiversity
868 distributions. - Nat. Commun. 6: 8221.
869 Needham, J. et al. 2018. Inferring forest fate from demographic data: from vital rates to
870 population dynamic models. - Proc. R. Soc. B Biol. Sci. 285: 20172050.
871 Normand, S. et al. 2014. Demography as the basis for understanding and predicting range
872 dynamics. - Ecography (Cop.). 37: 1149–1154.
873 Owens, H. L. et al. 2013. Constraints on interpretation of ecological niche models by limited
874 environmental ranges on calibration areas. - Ecol. Modell. in press.
875 Pagel, J. and Schurr, F. M. 2012. Forecasting species ranges by statistical estimation of
876 ecological niches and spatial population dynamics. - Glob. Ecol. Biogeogr. 21: 293–304.
877 Pagel, J. et al. 2020. Mismatches between demographic niches and geographic distributions are
878 strongest in poorly dispersed and highly persistent plant species. - Proc. Natl. Acad. Sci. U.
879 S. A. 117: 3663–3669.
880 Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the
881 distribution of species: are bioclimate envelope models useful? - Glob. Ecol. Biogeogr. 12:
882 361–371.
883 Pironon, S. et al. 2017. Geographic variation in genetic and demographic performance: new
884 insights from an old biogeographical paradigm. - Biol. Rev. 92: 1877–1909.
885 Quinn II, T. J. and Deriso, R. B. 1999. Quantitative fish dynamics. - Oxford University Press.
886 Quintana-Ascencio, P. F. et al. 2018. Predicting landscape-level distribution and abundance:
887 Integrating demography, fire, elevation and landscape habitat configuration (S Zhou, Ed.). -
888 J. Ecol. 106: 2395–2408.

889 Ramula, S. et al. 2008. General guidelines for invasive plant management based on comparative
890 demography of invasive and native plant populations. - J. Appl. Ecol. 45: 1124–1133.
891 Salguero-Gómez, R. et al. 2015. The compadre Plant Matrix Database: An open online
892 repository for plant demography. - J. Ecol. 103: 202–218.
893 Salguero-Gómez, R. et al. 2016. COMADRE: A global data base of animal demography. - J.
894 Anim. Ecol. 85: 371–384.
895 Schurr, F. M. et al. 2012. How to understand species' niches and range dynamics: a demographic
896 research agenda for biogeography. - J. Biogeogr. 39: 2146–2162.
897 Sexton, J. P. et al. 2009. Evolution and Ecology of Species Range Limits. - Annu. Rev. Ecol.
898 Evol. Syst. 40: 415–436.
899 Sheth, S. N. and Angert, A. L. 2018. Demographic compensation does not rescue populations at
900 a trailing range edge. - Proc. Natl. Acad. Sci. U. S. A. 115: 2413–2418.
901 Shinozaki, K. and Kira, T. 1956. Intraspecific competition among higher plants. VII. Logistic
902 theory of the C-D effect. - J. Inst. Polytech. Osaka Cy.: 35–72.
903 Stohlgren, T. J. et al. 2011. Bounding species distribution models. - Curr. Zool. 57
904 Strogatz, S. H. 1994. Nonlinear Dynamics and Chaos. - CRC press.
905 Teller, B. J. et al. 2016. Linking demography with drivers: climate and competition (J Metcalf,
906 Ed.). - Methods Ecol. Evol. 7: 171–183.
907 Thuiller, W. et al. 2014. Does probability of occurrence relate to population dynamics? -
908 Ecography (Cop.). 37: 1155–1166.
909 Vanderwel, M. C. et al. 2013. Climate-related variation in mortality and recruitment determine
910 regional forest-type distributions. - Glob. Ecol. Biogeogr. 22: 1192–1203.
911 Villellas, J. et al. 2015. Demographic compensation among populations: what is it, how does it

912 arise and what are its implications? (J Hille Ris Lambers, Ed.). - Ecol. Lett. 18: 1139–1152.
913 Williams, J. L. et al. 2011. Distance to stable stage distribution in plant populations and
914 implications for near-term population projections. - J. Ecol. 99: 1171–1178.
915 Wood, S. N. 2017. Generalized additive models: An introduction with R, second edition. - CRC
916 Press.
917 Zurell, D. et al. 2010. The virtual ecologist approach: simulating data and observers. - Oikos 119:
918 622–635.
919 Zurell, D. et al. 2016. Benchmarking novel approaches for modelling species' range dynamics. -
920 Glob. Chang. Biol. 22: 2651–2664.
921