Antecedent effect models as an exploratory rather than predictive tool to understand the response of plant populations to climate drivers

Aldo Compagnoni[1,2,3.], Dylan Childs[4], Tiffany M. Knight[1,2,3], Roberto Salguero-Gómez[6,7,8]

[1] Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108, Halle (Saale), Germany.
Deutscher Platz 5e, Leipzig, 04103, Germany

[2] Department of Community Ecology, Helmholtz Centre for Environmental Research–UFZ, 06120 Halle (Saale), Germany.

[3] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstrasse 4, Leipzig 04103, Germany.

[4] Department of Animal and Plant Sciences, University of Sheffield. Western Bank, Sheffield S10 2TN, UK.

[6] Department of Zoology, University of Oxford. 11a Mansfield Road, Oxford, OX1 3SZ, United Kingdom.

[7] Max Planck Institute for Demographic Research, Konrad Zuse Strasse 1, Rostock, DE-18057, Germany

[8] Centre of Excellence in Environmental Decisions. The University of Queensland, St Lucia QLD 4071, Australia

Number of tables: 0

Number of figures: 4

Word count: 5,982

Number of references: 67

Online appendices: 1

Abstract

1. Understanding or predicting the responses of natural populations to climate is an urgent task for ecologists. However, studies linking the temporal dynamics of populations to climate remain limited. In population ecology, studies typically assume that populations respond only to the climate of the most recent growing season or year. However, evidence also exists that certain organisms respond to the climate in a lagged-response fashion.
2. Antecedent effect models are a set of statistical tools that help researchers select the past climatic factors that best predict a response (*e.g.*, survival, reproduction). These models use the evidence provided by the data to select, for example, specific months, or time windows correlated with a response. Thus, antecedent effect models could improve our ability to perform exploratory analysis and to predict the effects of temporal climatic variation on population dynamics.
3. Here, we compare the predictive performance of antecedent effect models against simpler models. We fit three antecedent effect models: (1) weighted mean models (WMM), which weigh the importance of monthly anomalies based on a Gaussian curve, (2) stochastic antecedent models (SAM), which weigh the importance of monthly anomalies using a Dirichlet process, and (3) regularized regressions using the Finnish Horseshoe prior (FHM), which estimate a separate effect size for each monthly anomaly. We compare these approaches to a linear model using a yearly climatic predictor, and a null model with no predictors. We use high resolution demographic data from 34 plant species ranging between 7 and 36 years of length, which reflect the typical temporal grain employed to study the effects of climate on populations. We fit models to the overall asymptotic population growth rate (λ), and to its

underlying vital rates (survival, development, and reproduction) to examine the extent to which potential lagged effects occur.

4. Antecedent effect models do not consistently outperform the linear or null models. Average differences in performance are small among models, response variables, and across sample sizes. Surprisingly, ranked in order of decreasing predictive performance, the best approaches are: null model > linear model > FHM > SAM > WMMs. Development is on average easier to predict than survival and reproduction, and these have higher predictive power than the overall rate of growth of the population, λ.
5. *Synthesis*: In temporal datasets with limited sample size, antecedent effect models are better suited as exploratory tools. These models can help identify which single monthly anomalies drive population dynamics, while controlling the effect of noise on estimates. As such, antecedent effect models can be a useful tool for hypothesis generation, and to increase our understanding of the links between climate and plant population dynamics.

Keywords: Bayesian moving window, COMPADRE Plant Matrix Database, cross-validation, exploratory data analysis, forecasting, matrix population model, population growth rate, regularization, sliding window, vital rate.

Introduction

Understanding and predicting population dynamics has an increasingly important branch of ecological research. In population ecology, the last decades have witnessed considerable development and application of a myriad of statistical methods aimed at understanding and predicting fluctuations in population dynamics (Besbeas et al., 2002; Caswell, 2001; DeAngelis & Grimm, 2014; Ellner et al., 2016; Plard et al., 2019). Some of this research has been motivated by the need to understand how global change drives population dynamics (Adler & HilleRisLambers, 2008; Knape & de Valpine, 2011; Post, 2013). Increasingly, special emphasis is being devoted to understanding how climatic drivers such as temperature and precipitation affect demographic rates (henceforth "vital rates") such as survival, development (*i.e.,* changes in size during the life cycle), and reproduction (Ehrlén et al. 2016; Teller et al. 2016). The end goal of this ambitious research agenda is a comparative understanding of what makes species vulnerable to climate change. Progress in this understanding requires synthetic studies that link demographic rates to climatic drivers using standardized methods (Compagnoni et al., 2021).

Studies linking population dynamics to climate drivers typically have limited temporal extent, and thus often identify climatic predictors, usually at the annual scale, *a priori*. Specifically, most population studies are short-term (Salguero-Gómez et al., 2015), with very few studies lasting more than 20 years (Compagnoni et al., 2021). These low sample sizes have historically pushed investigators to consider annual or within-year (e.g., seasonal) climatic effects (reviewed in Evers, Knight, Inouye, Miller, Salguero-Gómez, et al., 2021). Presumably, this choice has practical reasons: first, monthly or annual precipitation and temperature data is widely available in gridded datasets (Daly et al., 1994; Karger & Zimmermann, 2018). Second, annual anomalies are an intuitive choice to model annual variation in vital rates, and seasonal anomalies can be justified by

physiological or behavioural considerations (Catchpole et al., 2000; White et al., 1999). However, there is no consensus on which season should affect vital rates the most. For example, in plants the climate of the growing season is often invoked as a key driver of population dynamics, but there is also evidence highlighting the importance of the dormant season (Kreyling, 2010), or of lagged effects from previous years (Hacket-Pain et al., 2018). Recent empirical studies provided support for both of these alternatives (Czachura & Miller, 2020; Evers, Knight, Inouye, Miller, Salguero-Gómez, et al., 2021; Tenhumberg et al., 2018).

Ideally, investigators need to simultaneously identify the timing and magnitude of climatic effects from their data. To address this need, ecologists have developed models that automatically select the timing most relevant to population dynamics (Ogle et al., 2015; van de Pol & Cockburn, 2011). Henceforth, we refer to these models as "antecedent effect models". Such automatic selection simultaneously addresses issues of multiple comparisons, testing a large number of alternative climatic predictors. These models infer specific time periods when abiotic factors are most influential to population dynamics, dividing the climate (or any other relevant factor) of a time period (*e.g.*, a year) into sub-periods of equal lengths – usually months. Such weighing can occur by estimating the importance of each monthly anomaly using weights, which can be constrained follow a functional form (van de Pol & Cockburn, 2011) or not (Ogle et al., 2015). Alternatively, the weighing can be performed by estimating a different effect size for each monthly climate anomaly, using statistical regularization or informative priors to minimize overfitting (Hoerl & Kennard, 1970). These models have potential as both exploratory and predictive tools. The automatic estimation of the timing and importance of climatic effects is a tool to generate hypotheses. Moreover, these models could increase predictive ability by reflecting the temporal processes that link climatic drivers to population dynamics.

The general performance of antecedent effect models to predict population dynamics across species and across datasets of realistic size remains unknown. Because antecedent effect models reflect our expectations on how climate affects populations, they should have relatively high predictive ability. However, because real world datasets have relatively low samples sizes, simpler models might still be best at prediction. Moreover, we do know yet know whether certain vital rates can be predicted more easily than others. In principle, we expect population growth rates should be harder to predict than vital rates. This expectation is rooted in the theory that, to minimize chances of extinction, population growth rates will evolve low temporal variation (Tuljapurkar & Orzack, 1980), while this is not necessarily true for the underlying vital rates (Hilde et al., 2020; Pfister, 1998). Here, we systematically evaluate the performance of antecedent effect models in predicting the out-of-sample variation of vital rates (*e.g.*, survival, development, and reproduction) as well as the overall population growth rate (λ) of 34 plant species. We compare and contrast the predictive performance of three types of antecedent effect models to that of null models and models based on annual climatic anomalies. We use monthly temperature and precipitation as climatic drivers, and a set of species from the COMPADRE Plant Matrix Database (Salguero-Gómez *et al.* 2015) that range between 7 and 36 years in duration. We evaluate the predictive performance of these five models to: (i) understand whether antecedent effect models have generally higher out-of-sample predictive performance than models based on mean annual climate anomalies, and null models using no climate predictor; (ii) evaluate the relative performance of different implementations of antecedent effect models; and (iii) compare whether models including climate have a higher predictive ability for vital rates than for population growth rates.

Methods

Data and response variables

To compare the predictive performance of antecedent effect models to that of simpler models, we used long-term time series of vital rate data obtained from matrix population models (MPMs, hereafter) contained in the COMPADRE Plant Matrix Database (Salguero-Gómez et al., 2015). This database contains MPMs of natural populations where individuals have been categorized into discrete categories and their vital rates (*e.g.*, survival, development, reproduction) tracked in discrete time steps, most often annually (Caswell, 2001). We selected datasets from COMPADRE with at least seven contiguous years (*i.e.*, six MPMs) for a single natural population. We chose seven years subjectively, to balance the number of species with enough temporal replication with statistical power. Moreover, this period of time does not represents an uncommon dataset, given that half of plant demographic studies exceed four years in duration (Römer et al., 2021; Salguero-Gómez et al., 2015). This criterion resulted in 26 studies from 35 species, an average length of 15 years (median = 10; range = 7-36), and an average number of two populations (median = 1, range=1-6). By combining our temporal and spatial replication provided we had a total of 1,057 MPMs. We used these data to calculate the expectation of survival, development, reproduction and the population growth rate (λ) associated with each annual MPM. By "development" we refer to changes in size during the life cycle of an individual: in the literature this demographic rate is often referred to ask "growth". These four year-specific metrics refer to a population at its stable stage distribution: the distribution of discrete stages which would result if the vital rates of each yearly MPM were kept constant. We calculated the vital rates normalizing by the stable stage distribution. We refer to the stable stage distribution because we do not have the observed stage distribution of the populations at each time step. We calculated survival, development, and

reproduction using standard methods (Franco & Silvertown, 2004), and λ as the dominant eigenvalue of each MPM (Caswell, 2001) with the R package popbio (Stubben & Milligan, 2007). Here, $\lambda >$ (or $<$) 1 corresponds to a population that is expected to grow (decline) under stationary conditions (Caswell, 2001).

We used gridded climatic data to generate the climatic predictors for our statistical models. Specifically, the climatic predictors in our models include mean monthly air temperature and precipitation at the location of each of the 77 examined populations. We obtained these data from CHELSA (Karger et al., 2017), or from the PRISM (Daly et al., 1994) datasets. We used PRISM data for North American populations sampled before 1979 and after 2013: 14 out of our 77 populations (Supplementary Online Materials Table S1), while we used CHELSA data for all other populations. Both CHELSA and PRISM provide gridded climatic data at a 30 arcsec (~1 km^2) resolution. They incorporate data on topographic effects on climate, which is useful whenever a population is located on mountainous terrain. Gridded datasets such as these are advantageous, in that they approximate weather station climate averages well (Behnke et al., 2016), while providing a standardized source of climatic data.

Our climatic predictors consisted of the monthly climatic anomalies observed in the 36 months preceding each vital rate observation (*i.e.*, preceding the end month of a given demographic projection interval). These anomalies were *z*-scores (mean = 0, SD = 1) with respect to the long-term monthly mean. We calculated long-term means exceeding the recommended minimum of 30 years of data (World Meteorological Organization, 2017). For most datasets, we calculated the long-term monthly mean during the 35-year period between 1979 and 2013 for CHELSA data. This period differed for species analyzed using PRISM data. For *Astragalus scaphoides*, this period extended to year 2014 (Tenhumberg *et al.* 2018). Finally, for the 10 datasets collected

before 1979, we use a 50-year long climatic mean whose last observation corresponds to the last census of the demographic data (Supplementary Online Materials Table S1).

Modeling overview

Using vital rates and log(λ) as response variables, we compared and contrasted the predictive performance of five models (Fig. 1): the (1) null model (NM), with no climatic covariates, the (2) climate summary model (CSM, Dalgleish et al., 2011), using yearly climate as predictor, and three types of antecedent effect models: the (3) weighted means models (WMM, van de Pol et al., 2016; van de Pol & Cockburn, 2011b), the (4) stochastic antecedent model (SAM, Ogle et al., 2015), and the (5) Finnish horseshoe prior model (FHM, Hastie et al., 2009). Models 1-5 are ranked by increasing flexibility, by progressively relaxing key assumptions (Also reviewed in Fig. 1). NMs are a "model of the mean" that do not account for the effect of climatic drivers. CSMs assume that the monthly anomalies observed within a specific period have all equal weight, and that these anomalies affect population dynamics in the same direction. Therefore, this model uses as predictor the mean of 12 monthly anomalies. WMMs weigh the contribution of the 12 monthly anomalies based on a Gaussian function. This model assumes that adjacent monthly anomalies have a similar – or correlated – effect on population dynamics. For example, monthly anomalies could decrease in weight depending on their distance to the demographic census (Figure 1). SAMs assume that the weights of adjacent monthly anomalies are not necessarily correlated. Like in WMMs, SAMs assume that monthly anomalies influence population dynamics in the same direction. SAMs weight monthly anomalies using coefficients drawn from a Dirichlet distribution (Figure 1; Ogle et al., 2015). Finally, FHM fit a separate coefficient for each of the 12 monthly anomalies. Regularized regressions are the most flexible among these five models: they assume

198 the effects of adjacent monthly anomalies are not necessarily correlated, and that monthly
199 anomalies can have either positive or negative effects. Note that FHMs were not developed to
200 model climatic effects (Piironen & Vehtari, 2017). However, regularization approaches such as
201 FHMs have been previously employed to estimate climatic effects on vital rate information (e.g.
202 Tredennick et al., 2016).

203 We fit data from three separate years to account for the potential of lagged climatic effects,
204 for which there is convincing evidence in case studies For example, for three populations we
205 analyze in this study, one of *Cryptantha flava* and two of *Astragalus scaphoides* (Supplementary
206 Online Materials

207 Table S1), previous studies showed evidence of lagged climatic effects (Evers, Knight,
208 Inouye, Miller, Salguero-Gómez, et al., 2021; Tenhumberg et al., 2018). All the aforementioned
209 models except the NMs and CSMs use 12 monthly anomalies as predictors. However, we re-fit
210 each model using climate data from three separate years: each year corresponding to a subset of
211 the 36 monthly anomalies preceding each demographic observation. We used the 12 months of the
212 year t ending with a demographic observation, and then we used monthly observations one year
213 (t-1), and two years (t-2) preceding each demographic observation. Hence, we fit a total of 13
214 models for each vital rate or estimation of population growth rate and climate variable
215 combination.

216 Below, we describe the equations of the models we tested assuming a normal response
217 variable, $\log(\lambda)$. To facilitate model comparison and model convergence, we used weakly
218 informative priors for the intercept and variance terms (Lemoine, 2019). We document the priors
219 for all models, and the prior predictive checks we ran to choose them, in the section of the

Supplementary Online Materials on prior predictive checks. Our null is a model of the mean, such that

$$y_{ti} \sim Normal(\hat{y}, \sigma) \quad (1a)$$

$$\hat{y} = \alpha, \quad (1b)$$

$$\alpha \sim Normal(0,0.5), \quad (1c)$$

$$\sigma \sim Gamma(0.01,0.01), \quad (1d)$$

where y_{ti} refers to a log(λ) value, t refers to the year of observation, i to a replicate within years (i exceeds one in datasets that are spatially replicated), $\hat{y}$ is the mean model prediction, and α is a model intercept, the mean. In the CSM, we model the response variable as a function of $\bar{x}_t$, the mean of the 12 monthly anomalies (z-scores) over year t, as

$$\hat{y}_t = \alpha + \beta \bar{x}_t, \quad (2a)$$

$$\beta \sim Normal(0,1), \quad (2b)$$

where β is the coefficient for the effect of $\bar{x}_t$, and $\hat{y}$ is the mean model prediction referred to year t.

The first antecedent effect model is the WMM, which models the response as a function of x_{tk}, the monthly temperature deviations for year t and month k:

$$\hat{y}_t = \alpha + \beta \sum_k w_k x_{kt} \quad (3a)$$

$$w_k = g_k / \sum_k g_k \quad (3b)$$

$$g_k = \frac{1}{\sigma_w \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma_w}\right)^2} \quad (3c)$$

$$\mu \sim Normal(6.5,12)[0,12], \quad (3d)$$

$$\sigma_w \sim Normal(1,12)[1,], \quad (3e)$$

where the 12 x_k monthly anomalies prior to the end of the relevant demographic year, are each weighted by w_k. In turn, w_k is a normalized vector of Gaussian weights, g_k, with mean μ and scale

σ_w. The square brackets in Eq. 3d and 3e show the limits beyond which the potential values of these parameters are truncated.

The SAM, first proposed by Ogle et al. (2015), adds flexibility to the WMM above,

$$\hat{y}_t = \alpha + \beta \sum_k w_k x_{kt}, \qquad (4a)$$

$$\boldsymbol{w} = \text{Dirichlet}\left(\overbrace{1,\ldots,1}^{n=12}\right), \qquad (4b)$$

where vector $\boldsymbol{w}$ contains 12 weights that sum to 1, and w_k is a single weight that refer to monthly anomaly x_{kt} observed in month k and year t. The added flexibility originates because the values in vector $\boldsymbol{w}$ do not follow a functional form (albeit they are correlated). Henceforth, we refer to the term $\sum_k w_k x_{kt}$, in both Eq. 3a and 4a, as the "antecedent" of year t.

Finally, the FHM is a multiple regression whose coefficients are shrunk via the "Finnish" horseshoe prior proposed by Piironen & Vehtari (2017). These models estimate a separate slope value, β_k, for each k month. The Finnish horseshoe prior prevents overfitting by shrinking estimates towards zero. This model is fit as follows:

$$\hat{y}_t = \alpha + \sum_k \beta_k x_{kt}, \qquad (5a)$$

$$\beta_k \sim Normal\left(0, \tau\tilde{\theta}_k\right), \qquad (5b)$$

$$\tilde{\theta}_k = \frac{c\theta_k}{\sqrt{c^2+\tau^2\theta_k^2}}, \qquad (5c)$$

$$\theta_k \sim HalfCauchy(0,0.1), \qquad (5d)$$

$$c^2 \sim InvGamma\left(\frac{v}{2}, \frac{v}{2}s^2\right), \qquad (5e)$$

$$\tau \sim HalfCauchy(0, \tau_0), \qquad (5f)$$

$$\tau_0 = \frac{m_0}{M-m_0}\frac{\sigma}{\sqrt{N}}. \qquad (5g)$$

Here, 12 β_k values depend on a normal distribution whose standard deviation is controlled by τ, the "global scale", and 12 $\tilde{\theta}_k$ "local scales". The global scale τ tends to shrink every single monthly estimate β_k toward zero. On the other hand, the $\tilde{\theta}_k$ values allow the β_k monthly estimates to become relatively large. In the FHM, the combination of τ and $\tilde{\theta}_k$ shrinks all β_k coefficients toward zero, while controlling the magnitude of β_k values for which there is evidence for a substantial departure from zero.

As stated above, we modified these models to fit survival, development, and reproduction data, which are non-normally distributed. These modifications are required for two reasons that are linked to the characteristics of the beta- (survival and development) and gamma (reproduction) distributions. First, survival and development represent transition probabilities bounded between zero and one, and reproduction a non-negative quantity bounded between 0 and positive infinity. These supports require, respectively, a logit and log link functions. Second, we adopted parameterizations that allowed us to model the mean of these variables. We illustrate our modifications using the CSM (Eq. 2) as an example. First, survival and development are rates that vary from 0 to 1, so we modeled these response variables using beta-regressions:

$$y_{ti} \sim \text{Beta}(a, b), \qquad (6a)$$

$$a = \hat{y}_t \varphi \qquad (6b)$$

$$b = (1 - \hat{y}_t)\varphi \qquad (6c)$$

$$\text{logit}(\hat{y}_t) = \alpha + \beta \bar{x}_t \qquad (6d)$$

$$\varphi \sim Gamma(1,1), \qquad (6e)$$

where $\hat{y}_t$ refers to the mean prediction of either survival or development in year t, φ represents the scale of the data, a and b are the shape parameters of the Beta distribution, and y_{ti} is a single survival or development value observed in year t and spatial replicate i.

We modeled reproduction as gamma distributed, in this case using a log link that ensures positive predictions,

$$y_{ti} \sim \text{Gamma}\left(\kappa, \frac{\kappa}{\hat{y}_t}\right), \qquad (7a)$$

$$\log(\hat{y}_t) = \alpha + \beta \bar{x}_t. \qquad (7b)$$

$$\kappa \sim Gamma(0.01, 0.01), \qquad (7c)$$

In Eq. 7, $\hat{y}_t$ refers to the mean prediction for fecundity values in year t, κ is the shape parameter of the gamma distribution, and $\kappa/\hat{y}_t$ is its scale calculated through the maximum likelihood estimator, and y_{ti} refers to a single fecundity value observed in year t and spatial replicate i.

Model fitting

We fit models using a Hamiltonian Monte Carlo sampler via the R package rstan (Stan Development Team 2018). We fit models running four parallel chains for 4,000 iterations, with a burn in of 1,000 iterations, and thinning every other iteration. Because we had 13 alternative models, four demographic responses (log(λ) and three vital rates), two climatic predictors, and 36 datasets, we fit a total of 3,744 models. Fitting these many models made it unfeasible to assess model convergence one model at a time. Therefore, we opted to assess issues occurring across model types, response variables, and dataset length. We assessed model convergence calculating the proportion of models where issues were indicated by i) more than 10 divergent transition, ii) Gelman and Rubin convergence diagnostic, $\hat{R}$, of at least one parameter exceeding 1.1, iii) at least one parameter having low effective sample size, and iv) at least one parameter having high Monte Carlo standard error. We provide the relevance and the rationale for choosing these convergence diagnostics in the Supplementary Online Materials on convergence diagnostics.

Finally, we further tested that our models represented the data well by performing model Bayesian checks. We checked our models by calculating the Bayesian p-values of posterior predictive checks. Bayesian p-values are more conservative, and should be preferred to, posterior predictive checks based on a loss functions such as sum of squared residuals (Conn et al., 2018). To perform a sampled posterior predictive check, we extracted one random sample from the joint posterior of each model to produce 1,000 simulated datasets. We obtained Bayesian p-values by calculating the proportion of times the sum of squared residuals (SSR) of the simulated data exceeded that of observed data. A proportion between 0.025 and 0.975 suggests the model reproduces the observed data well.

Model comparison

We compared model performance for each vital rate, climate predictor, and species combination. We performed comparisons based on the log pointwise predictive density (LPPD) (Bernardo & Smith, 2009; Hooten & Hobbs, 2015) calculated through a leave-one-year-out (LOYO) cross validation (e.g. Tredennick et al., 2016). A LOYO cross-validation creates a training set by leaving out an entire year of data rather than a single data point. The higher the value of LPPD of the held-out data, the better the predictive performance of the model. Our model selection focused on a LOYO cross-validation, because we aim at predicting year-level variation in demographic rates. We performed this procedure because in our spatially replicated datasets, the climate is strongly correlated among different sites. In particular, 16 of our 36 datasets are spatially replicated, and the correlation among site specific climatic anomalies is always above 0.7.

We compared model performance by plotting the difference in ΔLPPD, the difference in LPPD with the best model. The $\Delta LPPD_i$ of model *i* is calculated as LPPD*i* – max(LPPD). For each species, we plotted heat maps of the $\Delta LPPD_i$ values for each model, flagging the best three models. We then ordered heat maps by data length along the y-axis, and by the year of climate data (*t*, *t*-1, and *t*-2) on the x-axis.

We addressed our four objectives testing for differences among the relative support of our models using Pseudo-Bayesian Model Averaging. We calculated model weights using Pseudo-Bayesian Model Averaging for each combination of species, response variable (survival, progression, fecundity, and log(λ)), and predictor (temperature or precipitation). We calculated model weights following equations 8 and 9 in Yao et al. (2018). LPPD relative weights using the same formula used for AIC weights, so that

$$w_k = \frac{\exp(LPPD_k - \frac{1}{2}\mathrm{se}(\mathrm{LPPD_k}))}{\sum_{k=1}^{K} \exp(LPPD_k - \frac{1}{2}se(LPPD_k))} \qquad (8)$$

where *k* refers to each model, *se* is the standard error, and *K* is 13, the total number of models tested for each combination of species, response variable, and climatic predictor. We arcsine transformed these LPPD weights to use them as the response variable to address our four objectives. We used these weights to address our four objectives using graphs and average statistics. First, we plotted boxplots of model weights across all datasets, grouping them based on model types. This provided an evaluation of (i) whether antecedent effect models outperformed simpler ones, and (ii) which antecedent effect model had the best predictive performance. Finally, we plotted model weights of climate models grouping them by response variable, to verify whether (iii) climate is a better predictor of vital rates than population growth rates.

Case study

We presented a case study to show, in practice, the potential of antecedent effect models as a exploratory and predictive tool. First, we show how antecedent effect models identify monthly climatic drivers of demographic responses, and we discuss how these inferences can inform biological hypotheses. Second, we show the degree to which antecedent effect models improve model fit. Here, we present data from five populations of *Astragalus tyghensis,* an endemic legume plant with extremely restricted range found in arid lands of Western North America (Thorpe & Kaye, 2008). We present models fit on the fecundity of *Astragalus tyghensis* using precipitation anomalies one year prior to the last demographic observation (year *t*-1). This climatic predictor had high predictive power in our model comparisons (Fig. 2).

We used this case study to show for each model its fit to data, the effect size of climate, and the weight of monthly anomalies. First, for the CSM, WMM, and SAM, we plotted the average model predictions of 200 posterior samples against the actual data points. The climatic predictor is represented by $\bar{x}_t$ (Eq. 2a) for the CSM, and by the antecedent, $\sum_k w_k x_{kt}$ (Eq. 3a and 4a), for the WMM and SAM. These plots show the strength and the uncertainty in the modeled relationships. Second, we represent the effect sizes of climate, and their uncertainty, using kernel density estimates and Bayesian credible intervals. For the CSM, WMM, and SAM, we show the kernel density estimates of the β parameters (Eq. 2-4a). For the FHM, we show the 95% credible intervals of the the 12 β_k values (Eq. 5a). Third, we the show the relative importance of monthly anomalies in the WMM and SAM by directly plotting the weights of monthly anomalies (Eq. 3b and 4b) and their uncertainty. In the FHM, the relative importance in monthly anomalies is directly shown by the credible intervals of the 12 β_k values.

Results

Model fits

However, only SAMs encountered convergence problems in up to 10% of model fits, while in WMMs and FHMs convergence issues affected less than 1% of fitted models (Supplementary Online Materials: Fig. S4-5). This pattern of convergence issues was consistent across the four convergence diagnostics (Supplementary Online Materials: Fig. S4). Finally, convergence issues were determined by dataset length: none of the 12 models fit on datasets longer than 15 years presented convergence issues (Supplementary Online Materials: Fig. S5). In accordance with results on convergence issues, plotting $\hat{R}$ values against effective sample size shows that convergence issues occur only for a subset of SAMs (Supplementary Online Materials: Fig. S6).

Bayesian checks provided little evidence for model misspecification, with 55 over 988 models having p-values exceeding the 0.025 and 0.975 thresholds (Supplementary Online Materials: Fig S7). This figure reflects what we would expect to observe if all models were specified correctly. In particular, the upper and lower thresholds should be exceeded in 5% of cases, while in our case it was exceeded in 5.5% of cases.

Case study

The results of antecedent effect models were striking and biologically interpretable, but they did not substantially improve the ability to predict the fecundity of *Astragaluys tygensis*. All models show that precipitation is negatively associated with fecundity. According to the WMM, the climate preceding a demographic census tends to be the most important in driving fecundity. The SAM and FHM have more precise results, indicating that the most important climate windows are

located two (May) and six (January) months prior to the demographic census. Thus, antecedent effect models indicate that in this species, demography is simultaneously correlated to the precipitation of the growing and dormant season. However, despite such striking result, the model fits of antecedent effect models and the CSM are almost identical (Fig. 2). Hence, despite the predictive performance of antecedent effect models is slightly better in this case (Fig. 3), this is a quantitatively small difference.

Model comparison

Contrary to expectations, we found no evidence that antecedent effect models have higher predictive power than climate summary models. CSMs were clearly the best models including climate, but the weight of evidence was similar to that of NMs (Fig. 34, Supplementary Online Materials: Fig. S8). Among antecedent effect models, FHMs performed best, even if their mean LPPD relative weight was on average 0.01 lower than that of both NMs and CSMs (Supplementary Online Materials: Fig. S8). The temporal replication of the studies did not improve average model performance, but it decreased the likelihood of fitting models with very low LPPD relative weights (Supplementary Online Materials: Fig. S9). Finally, we found that models including climate tended to predict development better than the other response variable (Fig. 3-4; Supplementary Online Materials: Fig S10).

Discussion

Understanding how climate affects the vital rates (survival, development, reproduction) of natural populations is challenging because the replication of available datasets is typically small. Antecedent effects models are exploratory tools that can provide substantial help in hypothesizing the windows of time most strongly associated with vital rates variation. Our case study showcases a particularly clear case of the hypotheses that can be formulated with the aid of these models. With respect to the predictive ability of antecedent effect models, we drew three main findings from our comparison of LPPD weights. First, our original expectation that antecedent effect models would improve the ability to predict demographic responses to climate in common plant datasets was not upheld. Second, simple CSMs showed on average slightly better predictive ability than antecedent effect models. Third, climatic effects were revealed more often for development than for the other response variables. These findings suggest that, given current sample size limitations, our ability to predict the effect of climate on plant demographic rates is best attained with simple linear models. Antecedent effect models will become important for prediction only with large datasets, as suggested by results in other fields. Finally, our results might reflect the higher physiological sensitivity of development to relatively short-term climatic drivers. Capitalizing on this model comparison work, we conclude providing suggestions to practitioners interested in building these models in their own work.

Our case study exemplified both the exploratory power of antecedent effect models, and one potential reason for their poor predictive ability. In *Astragalus tyghensis*, both the SAM and FHM show a strong negative association between fecundity and the precipitation anomalies occurred in the May and January of the preceding year. These results raise at least three follow up hypotheses. First, such a strongly negative association with precipitation anomalies occurring a

438 year prior to a demographic observation could result from indirect effects mediated through
439 competition (Adler et al., 2012; Suttle et al., 2007) or physiological tradeoffs (e.g. costs of
440 reproduction, Crone et al., 2009; Miller et al., 2012). Second, the correlation with the precipitation
441 in January, a month when mean temperature is just above or below 0° Celsius, could reflect a
442 distinct effect of snowfall. The effects of snowfall on plants are often distinct from rain
443 (Compagnoni & Adler, 2014; Dalgleish et al., 2011); these effects could arise in different ways:
444 for example, because snow affects pathogens (Smull et al., 2019) or it protects plants from frost
445 (Inouye, 2000). Third, the large effect of May precipitation could reflect a key developmental
446 phase occurring either in May or June. All three hypotheses call for further tests to be performed
447 via analysis of existing data or field experiments.

448 Our case study led us to hypothesize that a reason why antecedent effect models do not
449 easily outperform the predictive ability of simpler models is correlations between monthly and
450 annual climate anomalies. In our case study, antecedent effect models do not visibly improve
451 model fit with respect to the CSM. However, the monthly precipitation anomaly with the strongest
452 effect, May, is strongly correlated (correlation = 0.5) with the annual precipitation anomaly. In
453 small datasets, such a strong correlation could make estimating the 2 to 12 additional parameters
454 associated with antecedent effect models disadvantageous for prediction.

455 The results of our model comparison indicate that simple linear models are an excellent
456 solution to maximize the ability of climatic drivers to predict plant demographic rates in small
457 datasets. This result is best understood in the context of the bias-variance tradeoff (Hastie et al.
458 2009). In small datasets, simple models will provide similar predictions even when fit on subsets
459 of the data. On the other hand, the variance in the predictions of more complex models can be
460 higher. Antecedent effect models control this variance by introducing bias in their estimates. The

461 best example are FHMs, which bias parameters towards zero (Piironen & Vehtari, 2017).
462 Similarly, the WMMs and SAMs introduce bias in the estimates of monthly weights via their
463 correlation. Thus, the weights referred to monthly anomalies (Eq. 3-4b) for which there is strong
464 evidence of an effect will draw down the remaining monthly weights. In our analysis, these biases
465 introduced by antecedent effect models were not sufficient to control the variance of their
466 predictions. Ecologists are typically interested in climatic predictors such as temperature and
467 precipitation because these are the focus of climate change projections (IPCC, 2013). However,
468 several other studies show that the predictive power of these climatic drivers on population
469 dynamics is often low (e.g. Harris et al., 2018; Knape & de Valpine, 2011; Tredennick et al., 2017).
470 Our findings suggest that this low predictive ability is general, and that it might reflect the
471 complexity of climatic effects on population dynamics. For example, process-based models of
472 plant growth, such as dynamic global vegetation models (Sitch et al., 2003) or certain crop models
473 (Jones et al., 2003), rely on a large number of specific climatic and soil predictors. Therefore, while
474 the assumptions of antecedent effect models are realistic, we hypothesize that improving
475 predictions in small datasets will require mechanistic models.

476 Given the slightly higher predictability of development with respect to other vital rates, we
477 cautiously hypothesize this might be a general pattern in plants. First, predicting log(λ) should be
478 harder, because species should evolve to buffer its variation through time (Hilde et al., 2020;
479 Pfister, 1998). Lower variation in λ will result in smaller effect sizes of climate, which then require
480 larger datasets to be detected statistically. Second, understanding how climate affects plant
481 mortality is a still a field in its infancy. For example, the survival of plants to drought is highly
482 specific to the species, plant functional type, and a large number of contextual factors (Volaire,
483 2018). Third, understanding the response of recruitment to climate is not straightforward: seed

484 production reflects indirect effects of development (Clark et al., 2021), and the transition from
485 seed to new recruits is a source of uncertainty (Paniw et al., 2017). On the other hand, the strong
486 dependence of development to climate is posited by physiology (Jones, 2013) and demonstrated
487 by ecosystem and crop models (Jones et al., 2003; Schulze et al., 2019).

488 Our model comparison suggests that the threshold of 20 years of data invoked in the
489 literature to predict the effects of climate on vital rates only holds for individual-level data, and
490 might be an optimistic generalization. This threshold was first indicated by a simulation study
491 using a functional smoothing method that produces estimates similar to our antecedent effect
492 models (Teller et al. 2016). This result was confirmed empirically employing the same functional
493 smoothing method on single 30-year (Hindle et al., 2019), and 26-year (Tenhumberg et al. 2018),
494 long datasets. However, single studies report successful results with smaller datasets. For example,
495 a study using functional smoothing methods on a 10-year long dataset (Scott et al., 2022), and a
496 study using a sliding-window approach (van de Pol et al., 2016) on two 15-years long individual
497 datasets (Evers et al. 2021). While these results are encouraging, we did not find a similar
498 predictive ability in our 12 datasets which were longer than 20 years. We suspect that our low
499 predictive ability could stem from using summary statistics rather than individual-level data. Our
500 datasets represent annual mean vital rates, while ignoring within-class individual-level variance in
501 response to climatic predictors. Hence, each of our data points could be biased by averaging across
502 size classes, micro-environments, and competitive environments. In addition, we are unaware of
503 whether the few studies to date that identify substantial climatic effects reflect selection bias. A
504 more solid test of these data replication thresholds implies a systematic study on a random sample
505 of individual-level datasets.

Our experience building and comparing antecedent effect models informs our suggestions on hyperparameter priors, and emphasizes the role of prior predictive checks. The choice of hyperparameter priors has strong effects on model fitting and performance. In WMMs, truncated hyperparameter priors (Eq. 3d-e) were key to allow model convergence by preventing the sampler from exploring unrealistic parameter values. In SAMs, we chose a vector of ones as the Dirichlet prior (Eq. 4b) to reach a balance between preventing overfitting, and allowing variation in the monthly weights. A substantially higher number for the prior (*e.g.*, a vector of 100s) would result in weights that are almost the same across months; a lower number would lead to overfitting. In FHMs, we suggest using a standard deviation of 0.1 for the prior on monthly effect sizes (Eq. 5d). The primary literature on FHMs usually employs a standard deviation of 1 (*e.g.*, Piironen & Vehtari, 2017), but in our prior predictive checks this choice generated unrealistically variable simulated data (Supplementary Online Materials: prior predictive checks). More generally, we found prior predictive checks were key as our initial prior choices produced simulated datasets with vastly different variance among model types and response variables. Priors generating unrealistically high variance had model convergence issues, and lower predictive performance (data not shown).

The complexity of the relationship between demography and climate and the small sample sizes available to researchers, make antecedent effect models an exploratory analysis, rather than predictive tool. Thus, antecedent effect models will help us sharpen our understanding of the relationships between climate and population dynamics. Such improved understanding will contribute to target future data collection efforts, and will advance mechanistic understanding. These exploratory analyses might be a required stepping stone for the maturation of forecasting in population ecology.

Acknowledgements

A.C. and TMK were supported by the Alexander von Humboldt Foundation under the framework of the Humboldt Professorship, and by the Helmholtz Recruitment Initiative of the Helmholtz Association (both to TMK); RSG was supported by NERC (R/142195-11-1). This work emanated from research discussions during the sAPROPOS (Analysis of PROjections of POpulationS) working group led by RSG at the German Centre for Integrative Biodiveristy Research (iDiv), which is funded by the German Research Foundation (FZT-118-202548816). We also thank the Max Planck Institute for Demographic Research for the support of the COMPADRE Plant Matrix Database and the hundreds of plant population ecologists who have contributed with demographic data to COMPADRE.

References

Adler, P. B., Dalgleish, H. J., & Ellner, S. P. (2012). Forecasting plant community impacts of climate variability and change: When do competitive interactions matter? *Journal of Ecology*, *100*, 478–487.

Adler, P. B., & HilleRisLambers, J. (2008). THE INFLUENCE OF CLIMATE AND SPECIES COMPOSITION ON THE POPULATION DYNAMICS OF TEN PRAIRIE FORBS. *Ecology*, *89*(11), 3049–3060. https://doi.org/10.1890/07-1569.1

Behnke, R., Vavrus, S., Allstadt, A., Albright, T., Thogmartin, W. E., & Radeloff, V. C. (2016). Evaluation of downscaled, gridded climate data for the conterminous United States. *Ecological Applications*, *26*(5), 1338–1351. https://doi.org/10.1002/15-1061

Bernardo, J. M., & Smith, A. F. M. (2009). *Bayesian Theory*. John Wiley & Sons.

Besbeas, P., Freeman, S. N., Morgan, B. J. T., & Catchpole, E. A. (2002). Integrating Mark–Recapture–Recovery and Census Data to Estimate Animal Abundance and Demographic Parameters. *Biometrics*, *58*(3), 540–547. https://doi.org/10.1111/j.0006-341X.2002.00540.x

Caswell, H. (2001). *Matrix population models. Construction, analysis, and interpretation.* Sinauer Associates.

Catchpole, E. A., Morgan, B. J. T., Coulson, T. N., Freeman, S. N., & Albon, S. D. (2000). Factors influencing Soay sheep survival. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *49*(4), 453–472. https://doi.org/10.1111/1467-9876.00205

Clark, J. S., Andrus, R., Aubry-Kientz, M., Bergeron, Y., Bogdziewicz, M., Bragg, D. C., Brockway, D., Cleavitt, N. L., Cohen, S., Courbaud, B., Daley, R., Das, A. J., Dietze, M., Fahey, T. J., Fer, I., Franklin, J. F., Gehring, C. A., Gilbert, G. S., Greenberg, C. H., … Zlotin, R. (2021). Continent-wide tree fecundity driven by indirect climate effects. *Nature Communications*, *12*(1), 1242. https://doi.org/10.1038/s41467-020-20836-3

Compagnoni, A., & Adler, P. B. (2014). Warming, soil moisture, and loss of snow increase Bromus tectorum's population growth rate. *Elementa: Science of the Anthropocene*, *2*, 000020. https://doi.org/10.12952/journal.elementa.000020

Compagnoni, A., Levin, S., Childs, D. Z., Harpole, S., Paniw, M., Römer, G., Burns, J. H., Che-Castaldo, J., Rüger, N., Kunstler, G., Bennett, J. M., Archer, C. R., Jones, O. R., Salguero-Gómez, R., & Knight, T. M. (2021). Herbaceous perennial plants with short generation time have stronger responses to climate anomalies than those with longer generation time. *Nature Communications*, *12*(1), 1824. https://doi.org/10.1038/s41467-021-21977-9

Conn, P. B., Johnson, D. S., Williams, P. J., Melin, S. R., & Hooten, M. B. (2018). A guide to Bayesian model checking for ecologists. *Ecological Monographs*, *88*(4), 526–542. https://doi.org/10.1002/ecm.1314

Crone, E. E., Miller, E., & Sala, A. (2009). How do plants know when other plants are flowering? Resource depletion, pollen limitation and mast-seeding in a perennial wildflower. *Ecology Letters*, *12*(11), 1119–1126. https://doi.org/10.1111/j.1461-0248.2009.01365.x

Czachura, K., & Miller, T. E. X. (2020). Demographic back-casting reveals that subtle dimensions of climate change have strong effects on population viability. *Journal of Ecology*, *108*(6), 2557–2570. https://doi.org/10.1111/1365-2745.13471

Dalgleish, H. J., Koons, D. N., Hooten, M. B., Moffet, C. A., & Adler, P. B. (2011). Climate influences the demography of three dominant sagebrush steppe plants. *Ecology*, *92*(1), 75–85. https://doi.org/10.1890/10-0780.1

Daly, C., Neilson, R. P., & Phillips, D. L. (1994). A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal of Applied Meteorology*, *33*(2), 140–158. https://doi.org/10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2

DeAngelis, D. L., & Grimm, V. (2014). Individual-based models in ecology after four decades. *F1000Prime Reports*, *6*. https://doi.org/10.12703/P6-39

Ehrlén, J., Morris, W. F., von Euler, T., & Dahlgren, J. P. (2016). Advancing environmentally explicit structured population models of plants. *Journal of Ecology*, *104*(2), 292–305. https://doi.org/10.1111/1365-2745.12523

Ellner, S. P., Childs, D. Z., & Rees, M. (2016). *Data-driven Modelling of Structured Populations: A Practical Guide to the Integral Projection Model*. Springer International Publishing. https://doi.org/10.1007/978-3-319-28893-2

Elton, C. S. (1924). Periodic Fluctuations in the Numbers of Animals: Their Causes and Effects. *Journal of Experimental Biology*, *2*(1), 119–163. https://doi.org/10.1242/jeb.2.1.119

Evers, S. M., Knight, T. M., Inouye, D. W., Miller, T. E. X., Salguero-Gómez, R., Iler, A. M., & Compagnoni, A. (2021). Lagged and dormant season climate better predict plant vital rates than climate during the growing season. *Global Change Biology*, *27*(9), 1927–1941. https://doi.org/10.1111/gcb.15519

Evers, S. M., Knight, T. M., Inouye, D. W., Miller, T. E. X., Salguero-Gómez, R., Iler, A. M., & Compagnoni, A. (2021). *Global Change Biology*, *n/a*(n/a). https://doi.org/10.1111/gcb.15519

Franco, M., & Silvertown, J. (2004). A Comparative Demography of Plants Based upon Elasticities of Vital Rates. *Ecology*, *85*(2), 531–538.

Fritts, H. (2012). *Tree Rings and Climate*. Elsevier.

Hacket-Pain, A. J., Ascoli, D., Vacchiano, G., Biondi, F., Cavin, L., Conedera, M., Drobyshev, I., Liñán, I. D., Friend, A. D., Grabner, M., Hartl, C., Kreyling, J., Lebourgeois, F., Levanič, T., Menzel, A., Maaten, E. van der, Maaten-Theunissen, M. van der, Muffler, L., Motta, R., … Zang, C. S. (2018). Climatically controlled reproduction drives interannual growth variability in a temperate tree species. *Ecology Letters*, *21*(12), 1833–1844. https://doi.org/10.1111/ele.13158

Harper, J. L. (1967). A Darwinian Approach to Plant Ecology. *Journal of Ecology*, *55*(2), 247–270. JSTOR. https://doi.org/10.2307/2257876

Harris, D. J., Taylor, S. D., & White, E. P. (2018). Forecasting biodiversity in breeding birds using best practices. *PeerJ*, *6*, e4278. https://doi.org/10.7717/peerj.4278

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer Science & Business Media.

Hilde, C. H., Gamelon, M., Sæther, B.-E., Gaillard, J.-M., Yoccoz, N. G., & Pélabon, C. (2020). The Demographic Buffering Hypothesis: Evidence and Challenges. *Trends in Ecology & Evolution*, *35*(6), 523–538. https://doi.org/10.1016/j.tree.2020.02.004

Hindle, B. J., Pilkington, J. G., Pemberton, J. M., & Childs, D. Z. (2019). Cumulative weather effects can impact across the whole life cycle. *Global Change Biology*, *25*(10), 3282–3293. https://doi.org/10.1111/gcb.14742

Hoerl, A. E., & Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, *12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634

Hooten, M. B., & Hobbs, N. T. (2015). A guide to Bayesian model selection for ecologists. *Ecological Monographs*, *85*(1), 3–28.

Inouye, D. w. (2000). The ecological and evolutionary significance of frost in the context of climate change. *Ecology Letters*, *3*(5), 457–463. https://doi.org/10.1046/j.1461-0248.2000.00165.x

IPCC. (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.

Jones, H. G. (2013). *Plants and Microclimate: A Quantitative Approach to Environmental Plant Physiology* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511845727

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., Wilkens, P. W., Singh, U., Gijsman, A. J., & Ritchie, J. T. (2003). The DSSAT cropping system model. *European Journal of Agronomy*, *18*(3), 235–265. https://doi.org/10.1016/S1161-0301(02)00107-7

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, *4*, 170122. https://doi.org/10.1038/sdata.2017.122

Karger, D. N., & Zimmermann, N. E. (2018). *CHELSAcruts – High resolution temperature and precipitation timeseries for the 20th century and beyond*. http://dx.doi.org/10.16904/envidat.159

Kaye, T. N., & Pyke, D. A. (2003). The Effect of Stochastic Technique on Estimates of Population Viability from Transition Matrix Models. *Ecology*, *84*(6), 1464–1476. https://doi.org/10.1890/0012-9658(2003)084[1464:TEOSTO]2.0.CO;2

Knape, J., & de Valpine, P. (2011). Effects of weather and climate on the dynamics of animal population time series. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1708), 985–992. https://doi.org/10.1098/rspb.2010.1333

Kreyling, J. (2010). Winter climate change: A critical factor for temperate vegetation performance. *Ecology*, *91*(7), 1939–1948. https://doi.org/10.1890/09-1160.1

Lemoine, N. P. (2019). Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, *128*(7), 912–928. https://doi.org/10.1111/oik.05985

Miller, T. E. X., Williams, J. L., Jongejans, E., Brys, R., & Jacquemyn, H. (2012). Evolutionary demography of iteroparous plants: Incorporating non-lethal costs of reproduction into integral projection models. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1739), 2831–2840. https://doi.org/10.1098/rspb.2012.0326

Ogle, K., Barber, J. J., Barron-Gafford, G. A., Bentley, L. P., Young, J. M., Huxman, T. E., Loik, M. E., & Tissue, D. T. (2015). Quantifying ecological memory in plant and ecosystem processes. *Ecology Letters*, *18*(3), 221–235. https://doi.org/10.1111/ele.12399

Paniw, M., Quintana-Ascencio, P. F., Ojeda, F., & Salguero-Gómez, R. (2017). Accounting for uncertainty in dormant life stages in stochastic demographic models. *Oikos*, *126*(6), 900–909. https://doi.org/10.1111/oik.03696

Pfister, C. A. (1998). Patterns of variance in stage-structured populations: Evolutionary predictions and ecological implications. *Proceedings of the National Academy of Sciences*, *95*(1), 213–218.

Piironen, J., & Vehtari, A. (2017). Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electronic Journal of Statistics*, *11*(2), 5018–5051. https://doi.org/10.1214/17-EJS1337SI

Plard, F., Turek, D., Grüebler, M. U., & Schaub, M. (2019). IPM ^{2}: Toward better understanding and forecasting of population dynamics. *Ecological Monographs*, e01364. https://doi.org/10.1002/ecm.1364

Post, E. (2013). *Ecology of Climate Change: The Importance of Biotic Interactions*. Princeton University Press.

Römer, G., Dahlgren, J. P., Salguero-Gómez, R., Stott, I. M., & Jones, O. R. (2021). *Plant demographic knowledge is biased towards short-term studies of temperate-region herbaceous perennials* (p. 2021.04.25.441327). bioRxiv. https://doi.org/10.1101/2021.04.25.441327

Salguero-Gómez, R., Jones, O. R., Archer, C. R., Buckley, Y. M., Che-Castaldo, J., Caswell, H., Hodgson, D., Scheuerlein, A., Conde, D. A., Brinks, E., de Buhr, H., Farack, C., Gottschalk, F., Hartmann, A., Henning, A., Hoppe, G., Römer, G., Runge, J., Ruoff, T., … Vaupel, J. W. (2015). The COMPADRE Plant Matrix Database: An open online repository for plant demography. *Journal of Ecology*, *103*(1), 202–218. https://doi.org/10.1111/1365-2745.12334

Schulze, E.-D., Beck, E., Buchmann, N., Clemens, S., Müller-Hohenstein, K., & Scherer-Lorenzen, M. (2019). Dynamic Global Vegetation Models. In E.-D. Schulze, E. Beck, N. Buchmann, S. Clemens, K. Müller-Hohenstein, & M. Scherer-Lorenzen (Eds.), *Plant Ecology* (pp. 843–863). Springer. https://doi.org/10.1007/978-3-662-56233-8_22

Scott, E. R., Uriarte, M., & Bruna, E. M. (2022). Delayed effects of climate on vital rates lead to demographic divergence in Amazonian forest fragments. *Global Change Biology*, *28*(2), 463–479. https://doi.org/10.1111/gcb.15900

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., & Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography

and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*(2), 161–185. https://doi.org/10.1046/j.1365-2486.2003.00569.x

Smull, D. M., Pendleton, N., Kleinhesselink, A. R., & Adler, P. B. (2019). Climate change, snow mold and the Bromus tectorum invasion: Mixed evidence for release from cold weather pathogens. *AoB PLANTS*, *11*(5), plz043. https://doi.org/10.1093/aobpla/plz043

Stubben, C., & Milligan, B. (2007). Estimating and Analyzing Demographic Models Using the popbio Package in R. *Journal of Statistical Software*, *22*, 1–23. https://doi.org/10.18637/jss.v022.i11

Suttle, K. B., Thomsen, M. A., & Power, M. E. (2007). Species Interactions Reverse Grassland Responses to Changing Climate. *Science*, *315*(5812), 640–642. https://doi.org/10.1126/science.1136401

Teller, B. J., Adler, P. B., Edwards, C. B., Hooker, G., & Ellner, S. P. (2016). Linking demography with drivers: Climate and competition. *Methods in Ecology and Evolution*, *7*(2), 171–183. https://doi.org/10.1111/2041-210X.12486

Tenhumberg, B., Crone, E. E., Ramula, S., & Tyre, A. J. (2018). Time-lagged effects of weather on plant demography: Drought and Astragalus scaphoides. *Ecology*, *99*(4), 915–925. https://doi.org/10.1002/ecy.2163

Thorpe, A. S., & Kaye, T. N. (2008). *Astragalus tyghensis: Actual vs. Predicted population sizes* (p. 18). Institute for Applied Ecology for Prineville District BLM.

Tredennick, A. T., Hooten, M. B., & Adler, P. B. (2016). Do we need demographic data to forecast plant population dynamics? *Methods in Ecology and Evolution*, *8*(5), 541–551. https://doi.org/10.1111/2041-210X.12686

Tredennick, A. T., Hooten, M. B., & Adler, P. B. (2017). Do we need demographic data to forecast plant population dynamics? *Methods in Ecology and Evolution*, *8*(5), 541–551. https://doi.org/10.1111/2041-210X.12686

Tuljapurkar, S. D., & Orzack, S. H. (1980). Population dynamics in variable environments. I. Long-run growth rates and extinction. *Theoretical Population Biology*, *18*(3), 314–342.

van de Pol, M., Bailey, L. D., McLean, N., Rijsdijk, L., Lawson, C. R., & Brouwer, L. (2016). Identifying the best climatic predictors in ecology and evolution. *Methods in Ecology and Evolution*, *7*(10), 1246–1257. https://doi.org/10.1111/2041-210X.12590

van de Pol, M., & Cockburn, A. (2011). Identifying the Critical Climatic Time Window That Affects Trait Expression. *The American Naturalist*, *177*(5), 698–707. https://doi.org/10.1086/659101

Volaire, F. (2018). A unified framework of plant adaptive strategies to drought: Crossing scales and disciplines. *Global Change Biology*, *24*(7), 2929–2938. https://doi.org/10.1111/gcb.14062

White, M. A., Running, S. W., & Thornton, P. E. (1999). The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest. *International Journal of Biometeorology*, *42*(3), 139–145. https://doi.org/10.1007/s004840050097

World Meteorological Organization. (2017). *WMO Guidelines on the Calculation of Climate Normals* (World Meteorological Organization).

Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. (2018). Using Stacking to Average Bayesian Predictive Distributions (with Discussion). *Bayesian Analysis*, *13*(3). https://doi.org/10.1214/17-BA1091

Figure legends

Figure 1. Antecedent effect models to estimate the importance or the effect of monthly climate anomalies on demographic outcomes. These plots represent hypothetical Bayesian estimates of the A) weight and B) effect sizes of 12 monthly climate anomalies observed before a demographic outcome. In both panels, the symbols represent estimate means, and the extremes of the error bars represent the 95% credible intervals. In the models of panel A), the effect size of climate depends on the relative weights of monthly anomalies. The CSM is a linear regression where every monthly anomaly has the same weight. The climatic predictor in this regression is thus given by the mean of 12 monthly anomalies. Hence, monthly weights are identical, and no error bars are shown. In the WMMs, monthly weights follow a Gaussian curve, while in SAMs the monthly weights are determined by a Dirichlet distribution. Finally, panel B) shows how FHM tends to produce effect sizes centered around zero – except for the few for which there is strong evidence.

Figure 2. Case study showing the fits of the climate summary and the three antecedent effect models to the fecundity data of the *Astragalus tyghensis*, using precipitation anomalies two years preceding the demographic census. In each plot, columns represent, respectively, model fit, climate effect sizes, and monthly weights. The rows refer to model type: respectively the climate summary model (CSM), weighted mean model (WMM), stochastic antecedent model (SAM), and Finnish horseshoe model (FHM). In the first column, we show model fits using bivariate plots where the y-axis represents the response variable, and the x-axis the precipitation predictor. The predictor is either an annual weather anomaly, or the "antecedent". The antecedent is the summation represented in the second term of Equations 3a and 4a. In these plots on model fit, the thick black line represents the fit of the average model, and the black dots represent the data points. The thick

grey lines represent 200 model fits generated by randomly selecting the intercept and slope of the Bayesian posterior of each model. Note that we do not show the model fit for the FHM model, because in this case each model has 12 predictor variables: one for each monthly anomaly. Moreover, note that for the WWMs and SAMs, the 'antecedent' changes depending on the monthly weights (third column): the antecedents we show refer to the mean monthly weights.

In the second column, we show the climatic effect sizes of each model. In CSMs, WMMs, and SAMs, the climatic effect size is a single parameter: β, or the "slope" of the regressions described in Equations 2-4a. We represent these effect sizes plotting the histogram (grey bars) and the kernel density estimates (red line) of the posteriors of β. In FHMs, the climatic effects correspond to 12 separate β values, one for each monthly anomaly. We represented these effect sizes showing the 2.5th, 50th, and 97.5th quantile of these 12 β posteriors. We show these quantiles using a point range, where dots represent the 50th quantile, and the vertical lines extend from the 2.5th to the 97.5th quantile.

We show climatic weights for the WMMs and SAMs, the only classes of models for which model weights are available. For WMMs, we show a line plot that connects the monthly weights, on the y-axis, across the 12 months, on the x-axis. We show 200 lines, each representing a random selection of the model posterior. For SAMs, we represent monthly weights showing the 2.5th, 50th, and 97.5th quantile of the posteriors for each of the 12 monthly weights. We show these quantiles of monthly weights using a point range, where dots represent the 50th quantile, and the vertical lines extend from the 2.5th to the 97.5th quantile.

Figure 3. For models using precipitation as predictor, tile plot of predictive power based on differences in log pointwise predictive density, ΔLPPD. Tiles with lighter colors denote models with better predictive ability; the best model have a ΔLPPD of zero. Symbols denote best (circles), second best (triangle) and third best (square) model. Each panel refers to a separate response variable: A) survival, B) growth, C) fecundity, and D) log(λ). Rows refer to each dataset, and columns to each model. We compare five types of models, each referring to equations 1 through 5, respectively, in the main text. We compare a total of 13 models, because we fit the four models including climatic predictors three times. We fit these models using the climate observations observed the first, second, and third year prior to a demographic outcome.

Figure 4. For models using temperature as predictor, tile plot of predictive power based on differences in log pointwise predictive density (ΔLPPD). The layout of the plot is the same of figure 2.

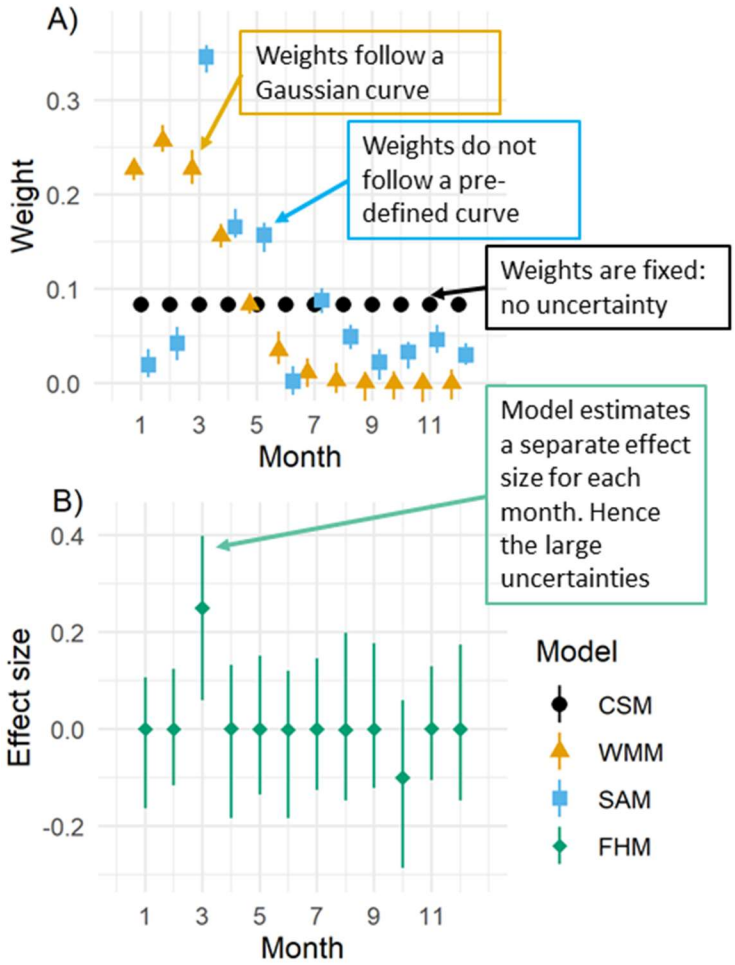
A)
Weights follow a Gaussian curve
Weights do not follow a pre-defined curve
Weights are fixed: no uncertainty
Weight
0.3
0.2
0.1
0.0
1
3
5
7
9
11
Month
B)
Model estimates a separate effect size for each month. Hence the large uncertainties
Effect size
0.4
0.2
0.0
-0.2
Model
CSM
WMM
SAM
FHM
1
3
5
7
9
11
Month

Figure 1

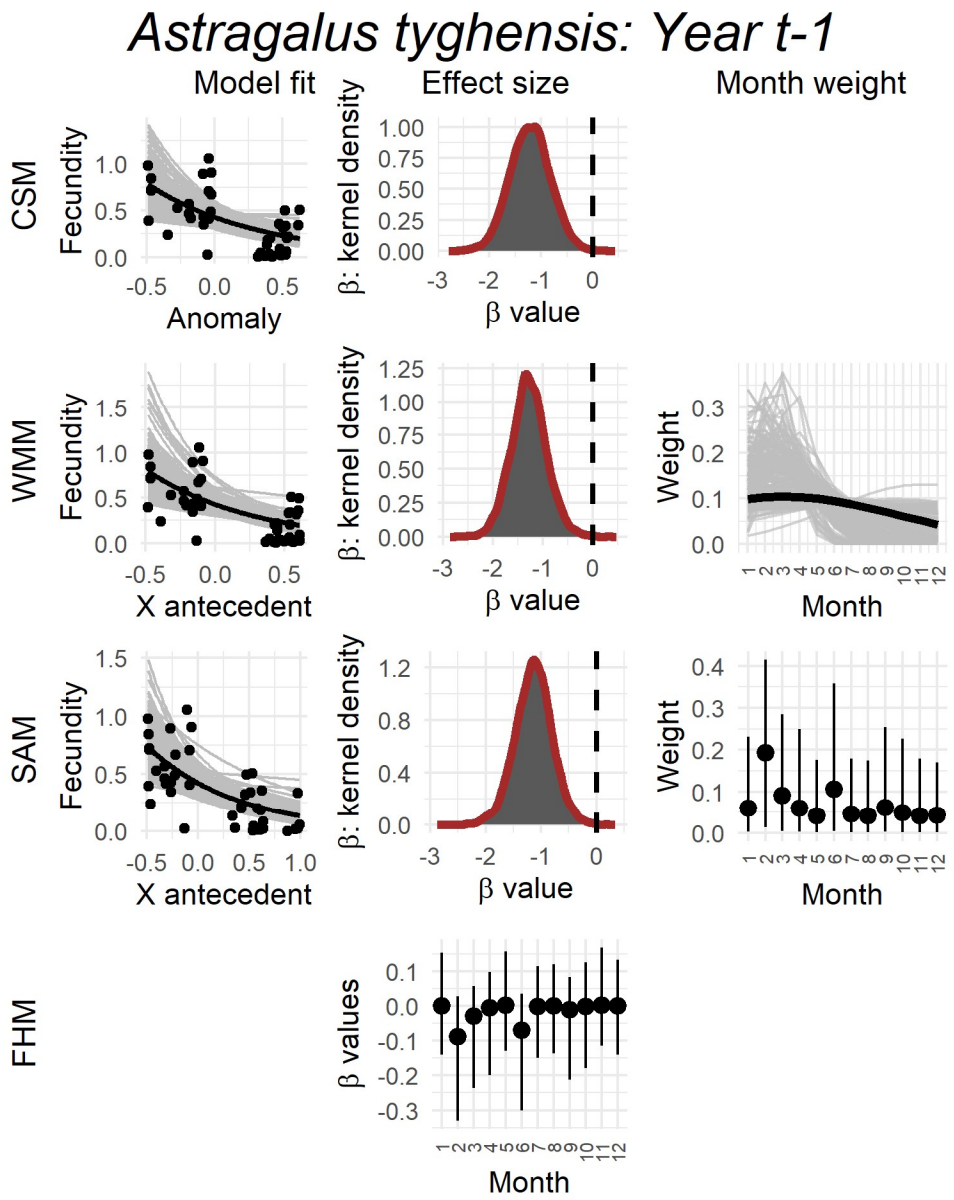
Astragalus tyghensis: Year t-1
Model fit
Effect size
Month weight
CSM
WMM
SAM
FHM
Fecundity
Anomaly
X antecedent
β: kernel density
β value
Weight
Month
β values

Figure 2

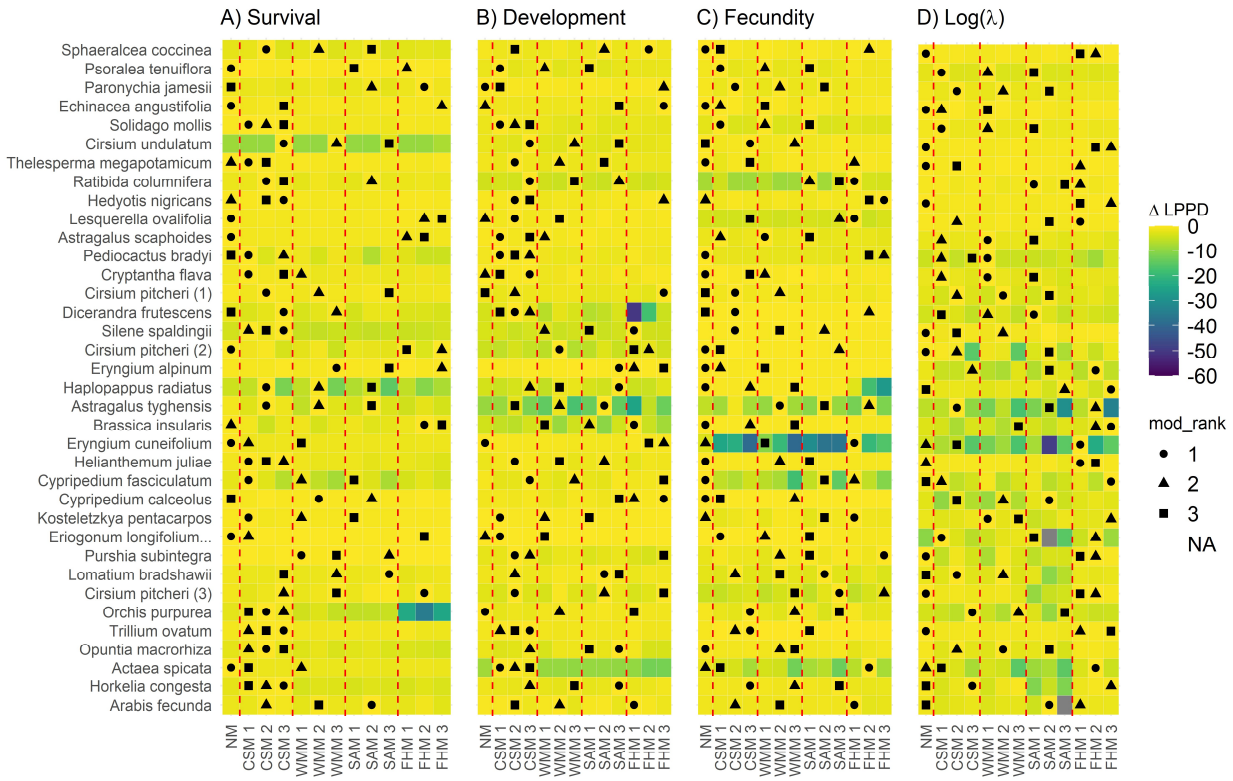
A) Survival
B) Development
C) Fecundity
D) Log(λ)
Sphaeralcea coccinea
Psoralea tenuiflora
Paronychia jamesii
Echinacea angustifolia
Solidago mollis
Cirsium undulatum
Thelesperma megapotamicum
Ratibida columnifera
Hedyotis nigricans
Lesquerella ovalifolia
Astragalus scaphoides
Pediocactus bradyi
Cryptantha flava
Cirsium pitcheri (1)
Dicerandra frutescens
Silene spaldingii
Cirsium pitcheri (2)
Eryngium alpinum
Haplopappus radiatus
Astragalus tyghensis
Brassica insularis
Eryngium cuneifolium
Helianthemum juliae
Cypripedium fasciculatum
Cypripedium calceolus
Kosteletzkya pentacarpos
Eriogonum longifolium...
Purshia subintegra
Lomatium bradshawii
Cirsium pitcheri (3)
Orchis purpurea
Trillium ovatum
Opuntia macrorhiza
Actaea spicata
Horkelia congesta
Arabis fecunda
NM
CSM 1
CSM 2
CSM 3
WMM 1
WMM 2
WMM 3
SAM 1
SAM 2
SAM 3
FHM 1
FHM 2
FHM 3
Δ LPPD
0
-10
-20
-30
-40
-50
-60
mod_rank
1
2
3
NA

Figure 3

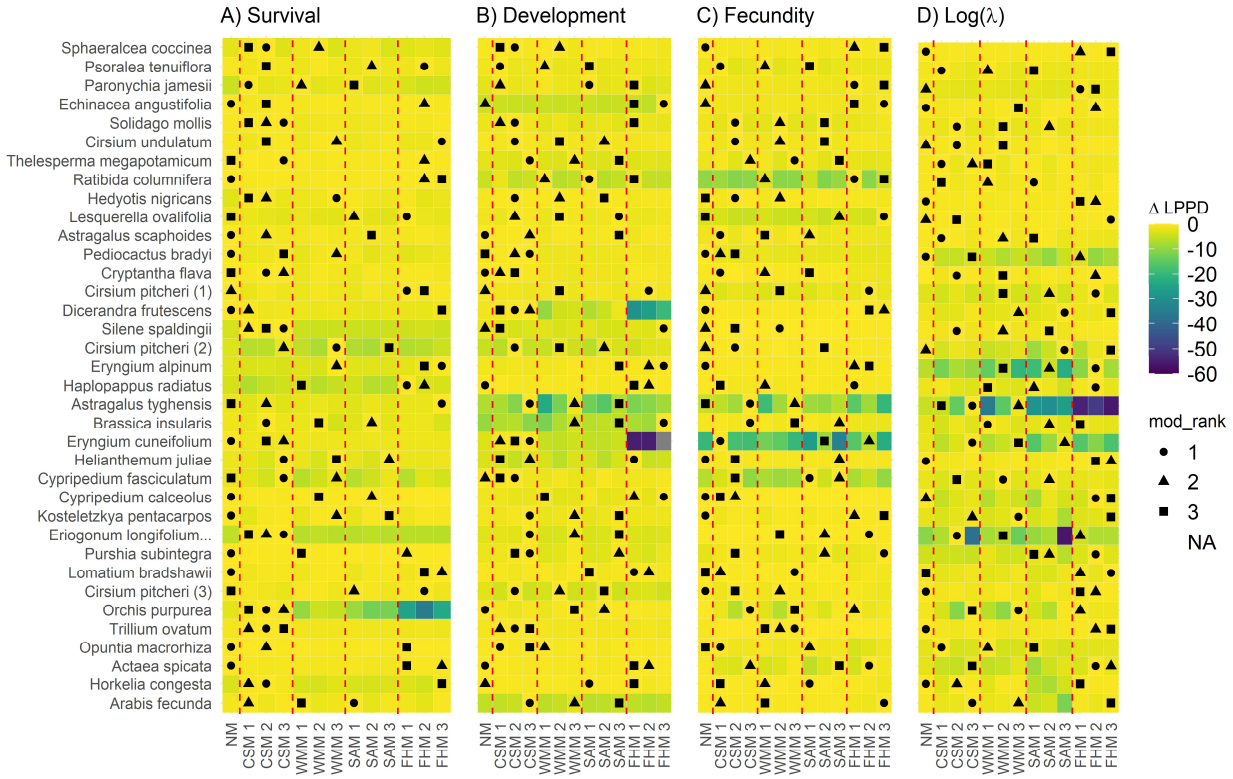
A) Survival
B) Development
C) Fecundity
D) Log(λ)
Sphaeralcea coccinea
Psoralea tenuiflora
Paronychia jamesii
Echinacea angustifolia
Solidago mollis
Cirsium undulatum
Thelesperma megapotamicum
Ratibida columnifera
Hedyotis nigricans
Lesquerella ovalifolia
Astragalus scaphoides
Pediocactus bradyi
Cryptantha flava
Cirsium pitcheri (1)
Dicerandra frutescens
Silene spaldingii
Cirsium pitcheri (2)
Eryngium alpinum
Haplopappus radiatus
Astragalus tyghensis
Brassica insularis
Eryngium cuneifolium
Helianthemum juliae
Cypripedium fasciculatum
Cypripedium calceolus
Kosteletzkya pentacarpos
Eriogonum longifolium...
Purshia subintegra
Lomatium bradshawii
Cirsium pitcheri (3)
Orchis purpurea
Trillium ovatum
Opuntia macrorhiza
Actaea spicata
Horkelia congesta
Arabis fecunda
NM
CSM 1
CSM 2
CSM 3
WMM 1
WMM 2
WMM 3
SAM 1
SAM 2
SAM 3
FHM 1
FHM 2
FHM 3
Δ LPPD
0
-10
-20
-30
-40
-50
-60
mod_rank
1
2
3
NA

Figure 4