Forecasting rodent population dynamics and community transitions with dynamic nonlinear models

Nicholas J. Clark[1,2], S. K. Morgan Ernest[3], Henry Senyondo[3], Juniper L. Simonis[3,4], Ethan P. White[3], Glenda M. Yenni[3], K. A. N. K. Karunarathna[1,2,5]

[1] School of Veterinary Science, Faculty of Science, The University of Queensland, Queensland 4343, Australia

[2] UQ Spatial Epidemiology Laboratory, School of Veterinary Science, The University of Queensland, Gatton, Queensland 4343, Australia

[3] Wildlife Ecology and Conservation, University of Florida, Gainesville, Florida 32611, USA

[4] DAPPER Stats, 3519 NE 15th Avenue, Suite 467, Portland, Oregon 97212, USA

[5] Department of Mathematics, Faculty of Science, Eastern University, Sri Lanka

Short title:

Modelling rodent community regime transitions

Corresponding author:

Nicholas J. Clark; n.clark@uq.edu.au

ABSTRACT

Ecological communities are dynamic. These dynamics are influenced by many sources of variation, making it difficult to understand or predict future change. Biotic interactions, and other sources of multi-species dependence, are major contributors. But ecological prediction overwhelmingly focuses on models that treat individual species in isolation. Here, we model the relative importance of nonlinear environmental responses and multi-species temporal dependencies for a community of semi-arid rodents. We use a hierarchical, Dynamic Generalized Additive Model (DGAM) to analyze monthly capture time series for nine rodents across a 25-year period. A vector autoregression to model unobserved trends allowed us to ask targeted questions about population dynamics. We find that multi-species dependencies are important for capturing unmeasured drivers of community change. Variation in captures for some species are expected to have delayed, often nonlinear effects on captures for others. These complexities are useful for inference but also for prediction. Models that captured multi-species dependence gave better near-term forecasts of community change than models that ignored it. We also quantify nonlinear effects of temperature change and positive effects of vegetation greenness on captures for nearly all species. Models that describe biological complexity, both through nonlinear covariate functions and multi-species dependence, are useful to ask targeted questions about population dynamics and drivers of change.

Keywords:

ecological forecasting; Generalized Additive Model; population dynamics; regime shift; time series; vector autoregression

INTRODUCTION

Regime shifts are sudden transitions from one ecological state to a stable, but different, state (Foley et al. 2003, Chaparro-Pedraza and de Roos 2020). Forecasting these events is a high priority (Biggs et al. 2009, Carpenter et al. 2011, Brook et al. 2013), and so it should be. Abrupt transitions can pose real threats to ecosystems and people that depend on them (Scheffer et al. 2009, Levin et al. 2013). Marine examples include catastrophic fishery collapses and shifts from kelp forests to sea urchin barrens (Roughgarden and Smith 1996, Biggs et al. 2009). On land, cattle grazing and fire suppression are linked to rapid transitions from semi-arid grassland to shrubland, magnifying erosion and driving declines of grassland species (Brandt et al. 2013, Cosentino et al. 2014). Other examples abound (May 1977, Anderson et al. 2008, Scheffer et al. 2009).

But most ecosystem changes are not sudden, they are gradual (Fukami and Wardle 2005, Hughes et al. 2013). To make better predictions, we must learn more about why gradual transitions happen. The relative abundances of species, for example, fluctuate for many reasons (Hampton et al. 2013). Food and shelter availability limit survival. Intraspecific (e.g. density dependence) and interspecific (e.g. competition, predation) interactions affect colonization and vital rates. Severe weather events and climate variation alter habitat suitability. Current changes in abundance can have carry-on effects on abundance in future time periods, irrespective of local conditions. These sources of variation combine to produce the observations we must analyze (Green et al. 2005, Greenville et al. 2016, Ovaskainen et al. 2017, Lasky et al. 2020). This makes it difficult to understand, let alone predict, ecosystem change.

In this paper, we pose a probability model to meet some of the challenges of inferring and predicting community dynamics. We use this model to analyse time series of long-term trapping data for multiple rodent species. Our work has two main goals. First, we make inferences about environmental and biotic factors that relate to population dynamics. Second, we forecast how the community is expected to change over short timescales. Our data come from the Portal Project, a

monitoring study designed to understand how rodent communities change over time (Brown 1998, Ernest et al. 2020). Connections between species interactions, environmental disturbance and population dynamics in this system have been extensively explored (Brown and Munger 1985, Heske et al. 1994, Brown 1998). Christensen et al. (2018) used latent time series clustering to infer that some rodent species respond synchronously to disturbances like droughts or storms. Using trait analyses, Supp et al. (2015) report that transient species are more diverse in their life history traits and show more chaotic spatial dynamics compared to established species, perhaps due to inferior competition. Manipulation experiments provide more direct evidence of biotic interactions. We know that some species can block others from arriving (Ernest and Brown 2001, Christensen et al. 2019a) or force them into less preferred habitat (Heske et al. 1994, Bledsoe and Ernest 2019). But do these complex interactions lead to regime transitions? If so, can we detect or predict them?

Unfortunately, we don't know if we can estimate, let alone forecast, community transitions in this system. Do species with similar requirements track environmental change in parallel? Or are their responses to change blurred by stochastic interactions? These questions are difficult. Previous work has made progress, but it has relied on pragmatic compromises. Brown and Heske (1990) used single species time series analyses and multi-species ordinations to detect patterns in rodent capture rates. But they did not use models that capture important statistical properties of the data, such as overdispersion and detection error. Christensen et al. (2018) inferred clusters using observed time series, but they could not infer drivers or generate species-level predictions. Simonis et al. (2021) and White et al. (2019) used time series models to forecast individual species. But they could not provide inferences about biotic effects. Bledsoe and Ernest (2019) used regression to relate the establishment of one species to the distribution of a competitor. Diaz and Ernest (2022) calculated energy use over time to make inferences about metacommunity dynamics.

All these studies were important and informative. But we still do not know how biotic and environmental relationships impact regime transitions. We address these gaps with a Bayesian

probability model that combines two powerful approaches. First, it uses nonlinear functions to estimate how species respond to environmental change. The model learns these functions hierarchically, where data for the community informs the response for each species. Second, our model estimates cross-dependences in species' capture rates. We show that decomposing time series analyses in this way allows us to ask more detailed questions about regime transitions.

MATERIALS AND METHODS

Rodent capture data

The Portal Project is a long-term study in the Chihuahuan Desert near Portal, Arizona (Ernest et al. 2020). In 1977 researchers began investigating rodent populations using baited trapping. A hierarchical sampling design includes 24 experimental plots (50m x 50m), each containing a grid of 49 Sherman traps (Brown 1998, Ernest et al. 2020). The design currently uses three treatments. In control plots (N = 10), holes in the fence are large enough to allow free access for all rodents. Full rodent removal plots (N = 6) have fences with no holes. Kangaroo rat exclosures (N = 8) have fences with small holes to allow passage of all rodents except the dominant kangaroo rats (*Dipodomys* genus). Investigators close holes during trapping to ensure all captured rodents are residents. Trapping follows the lunar cycle, with observations at approximately monthly intervals. Because of unforeseen interruptions, missing observations are common.

Open-source software exists to access Portal data (Christensen et al. 2019b, Simonis et al. 2022). We used these tools to extract trapping data (*portalR* version 3.134.0; downloaded October 2022). Our study focused on control plots from December 1996 – August 2022. The data has records for 20 rodent species, but some are rarely captured. We excluded species if they were observed in less than 10% of trapping sessions. We did this to focus inferences on species with the most influence on regime transitions. This left nine time series. Each observation was a vector of total

captures (pooled across all control plots) for the nine included species (Figure 1). On a small number of occasions (12 of 309 timepoints), two trapping sessions occurred in the same month. Because our models assumed one session per month, we retained the first session. This resulted in a loss of 3.9% of observations.

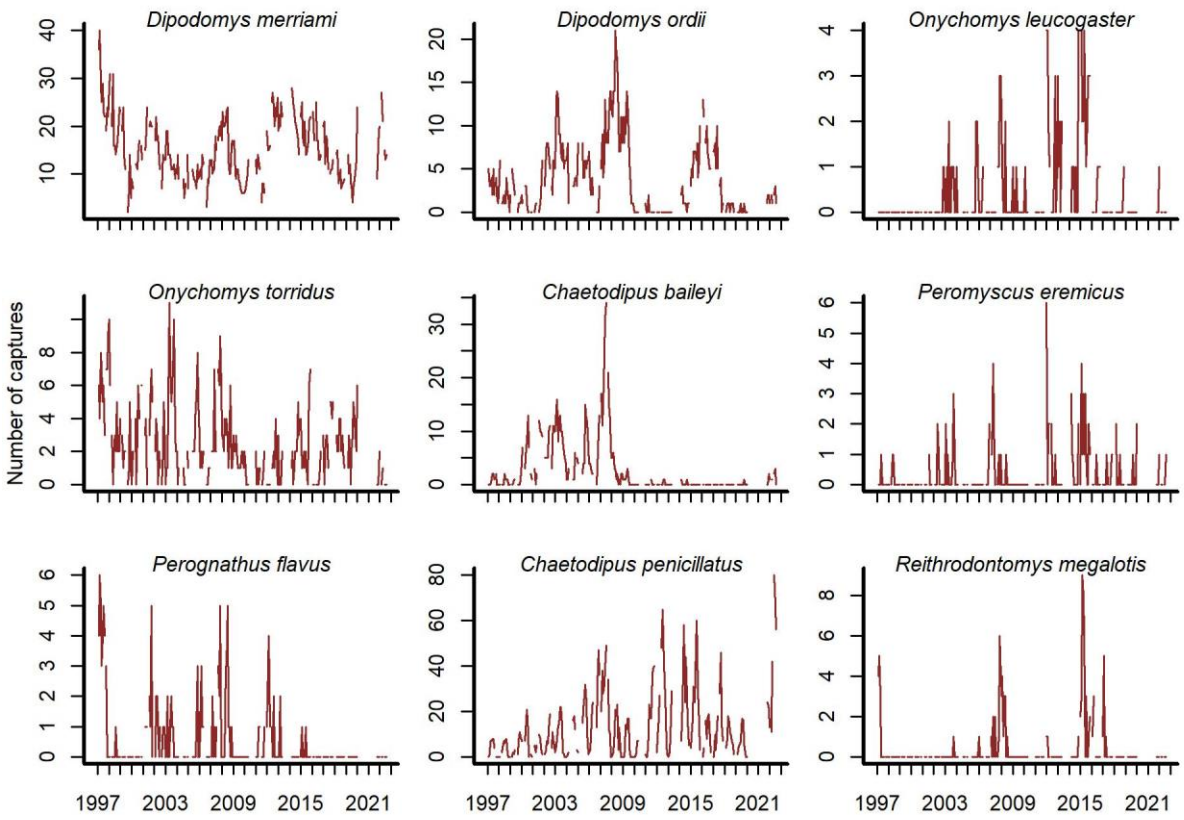


Figure 1: Rodent capture data from the Portal Project for the period December 1996 to August 2022. Counts are total captures across 10 control plots. Blanks are missing values.

Covariate measurements

Temperature and vegetation changes can impact rodent breeding and activity rates. They can also determine how much food and shelter is available (Cihlar et al. 1991, Pettorelli et al. 2011). We aimed to model these effects using minimum temperature and the Normalized Difference

Vegetation Index (NDVI) as covariates. Hourly air temperature (in °C) is recorded by an automated weather station. Landsat images are used to calculate NDVI (accessed from the United States Geological Service Earth Resources Observation and Science Center; https://www.usgs.gov/centers/eros). Measurements for both covariates were converted into monthly averages. We extracted covariate data from one year before the start of rodent captures (January 1995 to August 2022). This allowed us to calculate lagged and moving average versions.

A hierarchical Dynamic Generalized Additive Model

Hierarchical Generalized Linear Models (GLMs) capture contextual information in a grouped set of observations (Gelman 2006, McElreath and Koster 2014, McElreath 2020). They do this using adaptive regularization, where estimates for one group are informed by estimates from all other groups. This improves inferences and predictions, especially for noisy estimates. And noisy estimates are the norm in ecology. Other advantages of hierarchical models include reduced sensitivity to outliers and explicit group-level variance estimates (McElreath 2020). Provided computation is sound, a well-posed hierarchical model allows predictions to be generalized to unmeasured groups (Gelman 2006, McElreath and Koster 2014, Betancourt 2017). Many extensions are possible. For example, Pedersen and colleagues (2019) provide guidance for estimating hierarchical nonlinear functions in a Generalized Additive Model (GAM) framework. Hierarchical GLMs and GAMs can also capture overdispersion and autocorrelation, features that are common in ecology (Clark et al. 2001, Warton et al. 2015).

We used a hierarchical GAM to model rodent captures. We chose a GAM because species' responses to environmental change are expected to be nonlinear (Brown and Ernest 2002). Using a hierarchical model allowed us to jointly estimate these responses for nine species. There were several aspects of the data we needed to consider in our design. Monthly total rodent captures showed short-term fluctuations and long-term undulations (Supplementary Figure S1). Captures for

individual species also undulated over multi-annual cycles and were positively autocorrelated at lags up to 20 months (Supplementary Figure S2). We needed to model these temporal dynamics. The second aspect we needed to consider was multi-species dependence. Captures for some species were often synchronized with others. For example, the rise to dominance for the Bailey's pocket mouse (*C. baileyi*) in the late 1990's coincided with more captures for Ord's kangaroo rat (*D. ordii*) but fewer captures for other species (Ernest and Brown 2001). These included Merriam's kangaroo rat (*D. merriami*), Northern grasshopper mouse (*Onychomys leucogaster*) and Western harvest mouse (*Reithrodontomys megalotis*). When most species suffered declines during a drought in 2008 – 2009 (Christensen et al. 2018), the *C. baileyi* population crashed and has not recovered on control plots. This signaled a regime transition to a community dominated by *D. merriami* and the desert pocket mouse (*C. penicillatus*; Figure 1). Other notable aspects of the data were evidence of overdispersion and missing observations (Figure 1, Supplementary Figure S3). These features motivate techniques to deal with mean-variance relationships, missingness and multi-species dynamics. Our model tackled these challenges with the following form:

$$\boldsymbol{for\ i = 1, \ldots, 9\ rodent\ species}$$

$$\boldsymbol{for\ t = 1, \ldots, 309\ time\ steps}$$

$$\boldsymbol{Y_{i,t} \sim NegativeBinomial\left(\mu_{i,t}, \Phi_i\right)}$$

$$\boldsymbol{log\left(\mu_{i,t}\right) = \beta_{NDVI[i]} \cdot NDVI\_MA12_t +}$$

$$\boldsymbol{f_{global[t]}(Mintemp, lag) + f_{species[i,t]}(Mintemp, lag) + z_{i,t}}$$

Captures were modelled with a Negative Binomial observation process parameterized by location $\boldsymbol{\mu}$ and overdispersion $\boldsymbol{\Phi}$ (where $\mathbf{E}(\boldsymbol{Y}) = \boldsymbol{\mu}$ and $\mathbf{Var}(\boldsymbol{Y}) = \boldsymbol{\mu + \mu^2/\Phi}$). We determined priors for all parameters using prior simulations (Betancourt 2021). Our prior choices gave low probability to impossible outcomes while still allowing wide flexibility. One example was our prior for $\boldsymbol{\Phi}$. The Negative Binomial distribution generalizes to Poisson as $\boldsymbol{\Phi} \to \infty$. Conversely, small values of $\boldsymbol{\Phi}$ lead to extreme overdispserion that makes joint estimation of trends difficult (Clark and Wells 2022). We

sampled Φ from a containment prior using an $\boldsymbol{InverseGamma}$ distribution that placed high probability on large values ($\Phi < 600$) and low probability on small values ($\Phi < 2$):

$$\Phi \sim \boldsymbol{InverseGamma}(0.54, 6.88)$$

This complexity penalizing strategy pulled observations toward a Poisson distribution when support for overdispersion was lacking (Simpson et al. 2017).

Variation in $\boldsymbol{\mu}$ was modelled with a linear predictor capturing hierarchical functions of NDVI and minimum temperature. The structural forms of these functions were informed by theory and exploration of covariate time series (shown in Supplementary Figures S4-5). We used a 12-month moving average of NDVI ($\boldsymbol{NDVI_{MA12}}$) because we expected rodents to respond gradually to vegetation change. Our model assumed linear functions for effects of $\boldsymbol{NDVI_{MA12}}$, equivalent to a hierarchical slopes model. Slopes were drawn from a normal distribution with hyperparameters $\boldsymbol{\mu_{NDVI}}$ and $\boldsymbol{\sigma_{NDVI}}$. Because $\boldsymbol{NDVI_{MA12}}$was scaled to unit variance, we used a containment $\boldsymbol{InverseGamma}$ prior for $\boldsymbol{\sigma_{NDVI}}$ that placed low probability on small values that were not sensible and can cause computational challenges ($\boldsymbol{\sigma_{NDVI} < 0.1}$):

$$\boldsymbol{\beta_{NDVI}} \sim \boldsymbol{Normal}(\boldsymbol{\mu_{NDVI}}, \boldsymbol{\sigma_{NDVI}})$$
$$\boldsymbol{\mu_{NDVI}} \sim \boldsymbol{Normal}(0, 1)$$
$$\boldsymbol{\sigma_{NDVI}} \sim \boldsymbol{InverseGamma}(2.37, 0.73)$$

Autocorrelation functions for one species, desert pocket mouse (*Chaetodipus penicillatus*), showed cyclic patterns suggestive of seasonality (Supplementary Figure S3). This was expected, as the desert pocket mouse responds to falling temperatures and food shortages by entering a state of intermittent torpor (Brown and Zeng 1989). In the interest of parsimony, we could have used temperature effects to model seasonality only for this species, ignoring it for others. But autocorrelation functions can be misleading. They describe patterns in observations, and these

observations are usually a noisy representation of a latent process. We instead modeled seasonality for all species using functions of minimum temperature. But we did not believe species' captures would respond immediately to temperature change. Changing temperatures, and other signals of seasonality, induce physiological responses such as reproductive development (Kenagy and Bartholomew 1981). They also provide signals for plants to begin seed production. These responses take time. More realistic models allow for delayed responses to climate variation (Dickman et al. 1999, Luis et al. 2010, Wells et al. 2016). We used distributed lag nonlinear functions to capture seasonality through delayed effects of minimum temperature. These were estimated by finding penalized coefficients $\boldsymbol{\beta}$ for sets of basis functions $\boldsymbol{b}$. The $\boldsymbol{b}$ were constructed as tensor products of marginal basis functions from a cubic spline for lag (4 basis functions) and a thin plate spline for minimum temperature (basis functions). The tensor product enforced a spline in which functions of minimum temperature varied smoothly with increasing lag. To encourage multi-species learning, we included tensor products for the community $\boldsymbol{f_{global}}(\boldsymbol{Mintemp}, \boldsymbol{lag})$ and for species-level deviations $\boldsymbol{f_{species[i]}}(\boldsymbol{Mintemp}, \boldsymbol{lag})$:

$$\boldsymbol{f_{global}} = \sum \boldsymbol{\beta_{global}} \cdot \boldsymbol{b_{global}}$$

$$\boldsymbol{f_{species[i]}} = \sum \boldsymbol{\beta_{species[i]}} \cdot \boldsymbol{b_{species[i]}}$$

Basis coefficients $\boldsymbol{\beta}$ were given penalized multivariate normal priors. Prior precisions for these distributions were the products of penalty matrices $\boldsymbol{S}$ and regularization terms $\boldsymbol{\lambda}$:

$$\boldsymbol{\beta_{global}} \sim \boldsymbol{MultiNormal}(\mathbf{0}, \left(\sum \boldsymbol{S_{global}} \cdot \boldsymbol{\lambda_{global}}\right)^{-1})$$

$$\boldsymbol{\beta_{species[i]}} \sim \boldsymbol{MultiNormal}(\mathbf{0}, \left(\sum \boldsymbol{S_{species[i]}} \cdot \boldsymbol{\lambda_{species[i]}}\right)^{-1})$$

$$\boldsymbol{\lambda} \sim \mathbf{Normal}(\mathbf{30}, \mathbf{25})$$

The 'smoothing' parameters $\boldsymbol{\lambda}$ controlled function wiggliness by penalizing the second derivative between adjacent basis coefficients (Miller 2019, Pedersen et al. 2019). Penalty matrices $\boldsymbol{S}$ were constructed using basis expansion routines in the R package *mgcv* (Wood 2017). We used lags of up to six months in the past.

We refer to the hierarchical covariate effects (i.e. $\boldsymbol{\mu_{NDVI} + z_{NDVI[i]}\sigma_{NDVI} \cdot NDVI\_MA12_t + f_{global[t]}(Mintemp, lag) + f_{species[i,t]}(Mintemp, lag)}$) as the GAM component of the linear predictor. The remaining component, $\boldsymbol{z}$, used a multivariate dynamic model to capture lagged cross-dependencies. We used a vector autoregression (VAR) of order 1, where $\boldsymbol{z_t}$ was a 9-dimensional vector and $\boldsymbol{A}$ was a 9 x 9 matrix of autoregressive coefficients:

$$\boldsymbol{z_t \sim MultiNormal(A \cdot z_{t-1}, \Sigma_{VAR})}$$

Diagonal entries of $\boldsymbol{A}$ captured dependence of a species' trend (at time $\boldsymbol{t}$) on its own lagged values (at $\boldsymbol{t-1}$). Off-diagonals represented cross-dependencies. For example, the entry in $\boldsymbol{A}$**[2,3]** captured the effect of species $\boldsymbol{3}$'s trend at time $\boldsymbol{t-1}$ on the current trend for species $\boldsymbol{2}$ (at time $\boldsymbol{t}$). Conditional on these dependencies, latent trends were assumed to be contemporaneously independent. A diagonal covariance matrix $\boldsymbol{\Sigma_{VAR}}$ estimated trend variability. An informative $\boldsymbol{Beta}$ prior was used for standard deviations. Off-diagonals were fixed at zero:

$$\boldsymbol{A \sim Normal(0, 0.25)}$$
$$\boldsymbol{sqrt(\Sigma_{VAR[i,i]}) \sim Beta(8, 12)}$$
$$\boldsymbol{\Sigma_{VAR[i,i`]} = 0}$$

Benchmark models for forecast comparisons

Benchmarking against simpler models is useful for forecast evaluation (Simonis et al. 2021, Lewis et al. 2022). It is difficult to know if we are learning more about a system if our complex models cannot

produce better, or at least different, predictions than simple models. We refer to our model above as ***GAM-VAR*** for comparison against two benchmarks. The first used the same GAM linear predictor (and priors) as the ***GAM-VAR*** model, but replaced the VAR(1) with AR(1) trends:

$$z_{i,t} \sim Normal(AR_{i,t-1} \cdot z_{i,t}, \sigma_{AR[i]})$$

$$\sigma_{AR} \sim Beta(8, 12)$$

This model (called ***GAM-AR***) eliminated cross-dependencies among species' trends and was a natural simplification of ***GAM-VAR***. The second benchmark, referred to as ***AR***, also used AR(1) trends but removed the GAM component. Because this model only learned from past observations, comparisons against it helped us understand how covariates impacted predictions and inferences.

Estimation and forecast evaluation

We estimated posterior distributions for all models with Hamiltonian Monte Carlo in Stan (Carpenter et al. 2017, Stan Development Team 2022), specifically the *CmdStanr* interface (Gabry and Češnovar 2021). Stan's sampling algorithms provide state-of-the-art diagnostics for probabilistic models (Betancourt 2017). For example, Hamiltonian Markov chains diverged when attempting to estimate minimum temperature deviations for some species. Our data were not informative enough to learn how, or even if, these species responded to temperature change in ways that differed from the community response. Stan's diagnostics guided us to a model that could be reliably estimated, which included deviation functions for the four most frequently captured species (*Dipodomys ordii*, *D. merriami*, *Onychomys torridus* and *Chaetodipus penicillatus*). Posterior distributions were processed in R 4.2.0 (R Core Team 2020) with the *mvgam* R package (Clark and Wells 2022). Traceplots, rank normalized split-$\hat{R}$ (Vehtari et al. 2021) and effective sample sizes interrogated convergence of four parallel chains. Each chain was run for 500 warmup and 1000 sampling iterations.

Models were trained on the first 289 timepoints (~24 years). Remaining 20 timepoints were held out to evaluate forecasts. Although the total number of captures was usually forecast to be less than the number of traps (196 traps per session), on rare occasions our predicted total exceeded this number. But trapping more animals than we have traps available is unlikely (the observed total never exceeded 109). This conflict with domain knowledge led us to use judgmental forecast adjustment. We re-scaled any posterior draws where the total exceeded 196 using the following equation: $\widehat{\boldsymbol{Y}}_{\boldsymbol{i,t}} = \left(\widehat{\boldsymbol{Y}}_{\boldsymbol{i,t}} / \sum \widehat{\boldsymbol{Y}}_{\boldsymbol{t}}\right) \cdot \mathbf{196}$. Rescaling only affected a small percentage of posterior draws (0.40% of draws were rescaled in the ***GAM-VAR*** model). We then proceeded with forecast evaluation, which for probabilistic models should use the full distribution to better understand deficiencies (Simonis et al. 2021). We used the variogram score for evaluation. This proper scoring rule penalizes distributions that are less precise and that do not adequately capture observed correlations in test observations (Scheuerer and Hamill 2015). For completeness, we repeated evaluations using unscaled forecasts. R code to replicate all analyses and produce Figures is included in the Supplementary materials and will be permanently archived on Zenodo on acceptance.

RESULTS

All models showed adequate convergence and efficient posterior exploration. Rank normalized split-$\hat{R}$ values were <1.05 for all parameters and effective sample sizes ranged from 285 – 2952 (median = 1980). Because our models were complex, we did not rely on point estimates / intervals to interpret them. Instead, we use posterior predictive simulations to interrogate and compare models.

NDVI and minimum temperature effects

Our model consistently recovered environmental and climatic signals. Hierarchical $\boldsymbol{NDVI_{MA12}}$ slopes were positive with high probability for eight of nine species. Using the sample-average minimum

temperature and ignoring the trend, we simulated expected captures, $\mathbf{exp}(\boldsymbol{\mu})$, under different $\boldsymbol{NDVI_{MA12}}$ scenarios. The first resembled a relatively dry/brown vegetation state ($\boldsymbol{NDVI_{MA12}} = -\mathbf{0.75}$). The second resembled a relatively moist/green vegetation state ($\boldsymbol{NDVI_{MA12}} = \mathbf{0.75}$). We used 1,000 simulations for each scenario. The model gave higher probability to increased captures in the moist/green scenario for all species, but uncertainties varied (Figure 2). Greatest increases were expected for Ord's kangaroo rat (*D. ordii*), Western harvest mouse (*R. megalotis*) and cactus mouse (*Peromyscus eremicus*). The two species that have dominated control plots recently (*D. merriami* and *C. penicillatus*) showed relatively weak increases. The model was less confident about the direction of effect for Northern grasshopper mouse (*O. leucogaster*). For this species, the model expected an increase in 64% of simulations and a decrease in 36% (Figure 2).

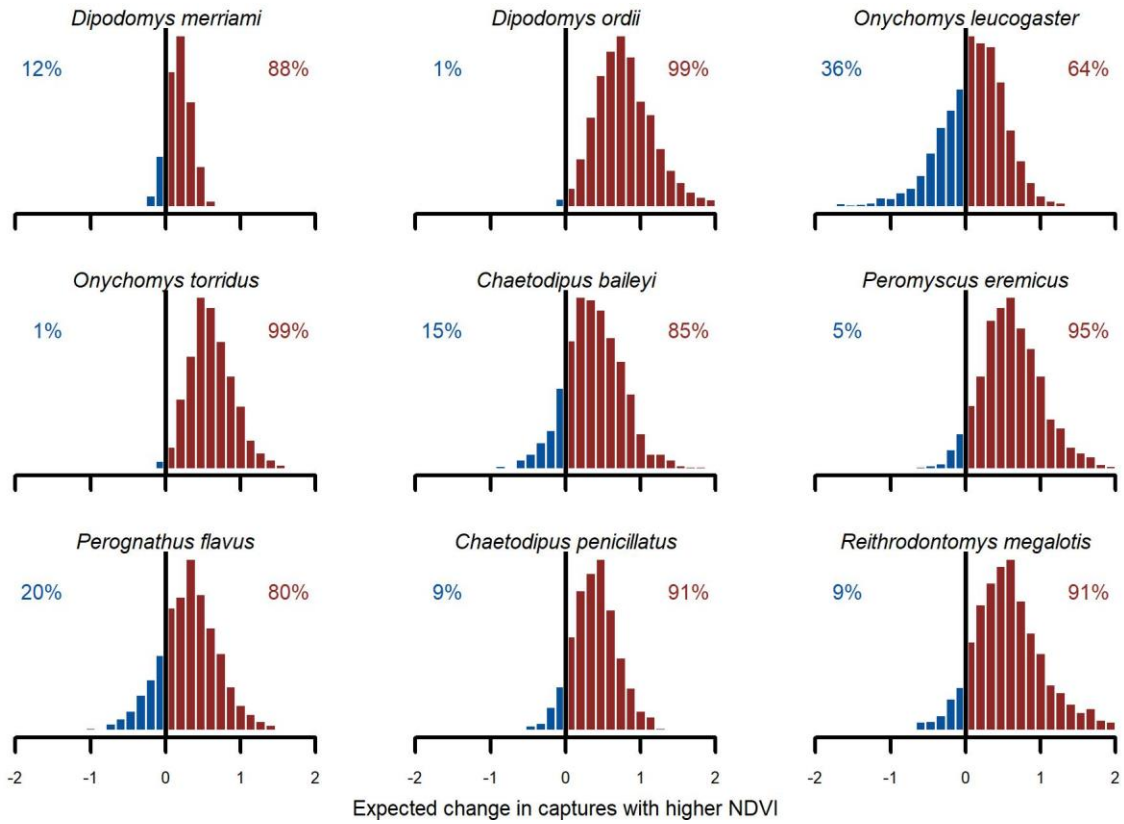


Figure 2: Posterior NDVI contrasts. Histograms illustrate how much the expected number of captures, $\mathbf{exp}(\boldsymbol{\mu})$, would change if the z-scored NDVI 12-month moving average ($\boldsymbol{NDVI_{MA12}}$)

316 changed from a relatively low value (-0.75) to a relatively high value (0.75). Numbers in each plot
317 indicate the proportion of probability mass at or below zero (in blue) vs above zero (in red).

318

319 Interpreting minimum temperature distributed lag effects also required simulation. We used
320 temperatures from 1997 to make posterior predictions. Multiplying a draw of coefficients with the
321 tensor product basis functions for 1997 gave us a realization from the set of posterior functions. We
322 visualized 1,000 simulated functions for each species (Figure 3). There was large uncertainty in
323 function shapes for all species except desert pocket mouse (*C. penicillatus*). Captures for this species
324 were expected to increase from May to October and decrease sharply in winter. For seven of the
325 other eight species, the model generally expected more captures in spring (March – May) and fewer
326 in late summer / autumn (July – October). But the shapes of these responses varied. The five species
327 that relied solely on the global function (*O. leucogaster*, *C. baileyi*, *P. eremicus*, *P. flavus* and *R.*
328 *megalotis*) were expected to show tighter spring peaks and autumn troughs. The two kangaroo rat
329 species (*D. merriami* and *D. ordii*) had smoother shapes that decreased gradually from mid-summer
330 to mid-winter. But the model expected *D. ordii* captures to peak slightly later (May as opposed to
331 March for *D. merriami*). The Southern grasshopper mouse (*O. torridus*) was the only species that was
332 expected to show higher captures in late autumn / early winter (Figure 3).

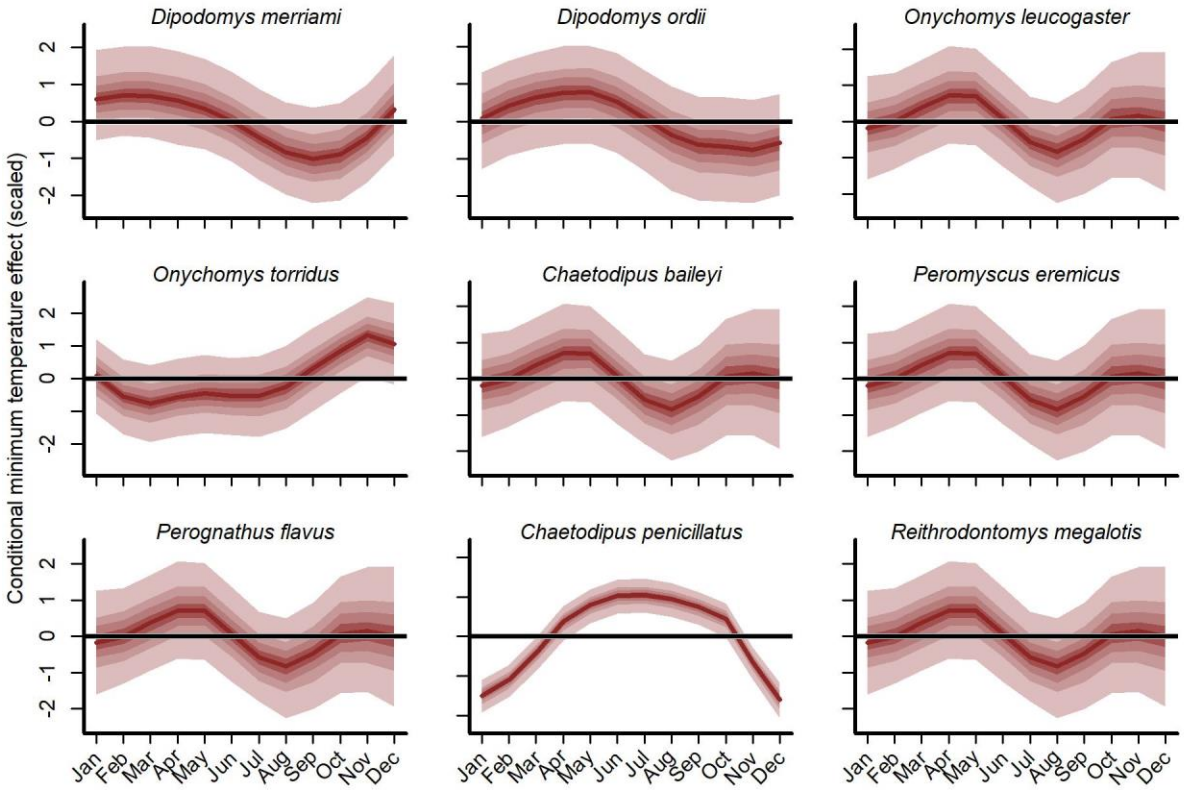


Figure 3: Conditional distributed lag minimum temperature functions, using temperatures observed in 1997. All other effects were ignored. Functions for *O. leucogaster*, *C. baileyi*, *P. eremicus*, *P. flavus* and *R. megalotis* were drawn solely from the global function. Functions for other species were the sum of the global function and a species-specific deviation function. Estimates were scaled to unit variance for better comparisons. Ribbon shading shows posterior empirical quantiles (90th, 60th, 40th and 20th). Dark red lines show posterior medians.

Cross-dependencies in latent trends

Posterior VAR(1) coefficients supported complex temporal dependencies. Autoregressive coefficients, which capture self-dependence, were large and positive for all species except the cactus mouse, *P. eremicus* (diagonal entries in Supplementary Figure S6). The model also estimated cross-dependence effects, where one species' trend was associated with variation in another's trend at the next timestep (off-diagonal entries in Supplementary Figure S6). Although many of these effects

were uncertain and centred on zero, there were prominent patterns. For example, the model expected fewer silky pocket mouse (*P. flavus*) captures if there was an increase in Ord's kangaroo rat (*D. ordii*) captures a month earlier (row seven, column two in Supplementary Figure S6).

It is tempting to walk through pairwise VAR effects one by one. But these coefficients only provide marginal insights into a network of conditional associations. It is better to interpret them jointly. We again used simulations, this time in the form of impulse response functions (Lütkepohl 1990). We generated a sudden 'impulse' in captures for one species and asked how captures for other species might change over the next six months. Responses to an impulse of three excess captures for Ord's kangaroo rat (*D. ordii*) are shown in Figure 4. Following a *D. ordii* pulse, the model expected fewer captures for six of the other eight species. But the shapes of these declines differed. Captures for silky pocket mouse (*P. flavus*), cactus mouse (*P. eremicus*), Merriam's kangaroo rat (*D. merriami*), Northern grasshopper mouse (*O. leucogaster*) and Western harvest mouse (*R. megalotis*) gradually declined in most simulations (Figure 4). The effect on desert pocket mouse (*C. penicillatus*) was also negative but decayed more quickly (Figure 4). These simulated responses were not only insightful, they also illustrated why VAR effects should not be interpreted in isolation. For example, the cross-dependence effect of *D. ordii* on *R. megalotis* was marginally positive (row nine, column two in Supplementary Figure S6). Yet the model expected a slow decline in *R. megalotis* captures after a *D. ordii* pulse. How could this be? We must examine all VAR effects for an explanation. If captures for *D. ordii* suddenly increased, the model expected fewer subsequent captures for *P. eremicus* and *O. leucogaster*. Both species were then expected to induce a decline for *R. megalotis* through their positive cross-dependence coefficients. This effect was magnified because *D. ordii*'s trend was autocorrelated. A *D. ordii* pulse should lead to elevated captures for this species over a few months. Different effects were expected when changing the focal species. For example, the model expected several species to show increased captures after a pulse of desert pocket mouse (*C. penicillatus*; Supplementary Figure S7).

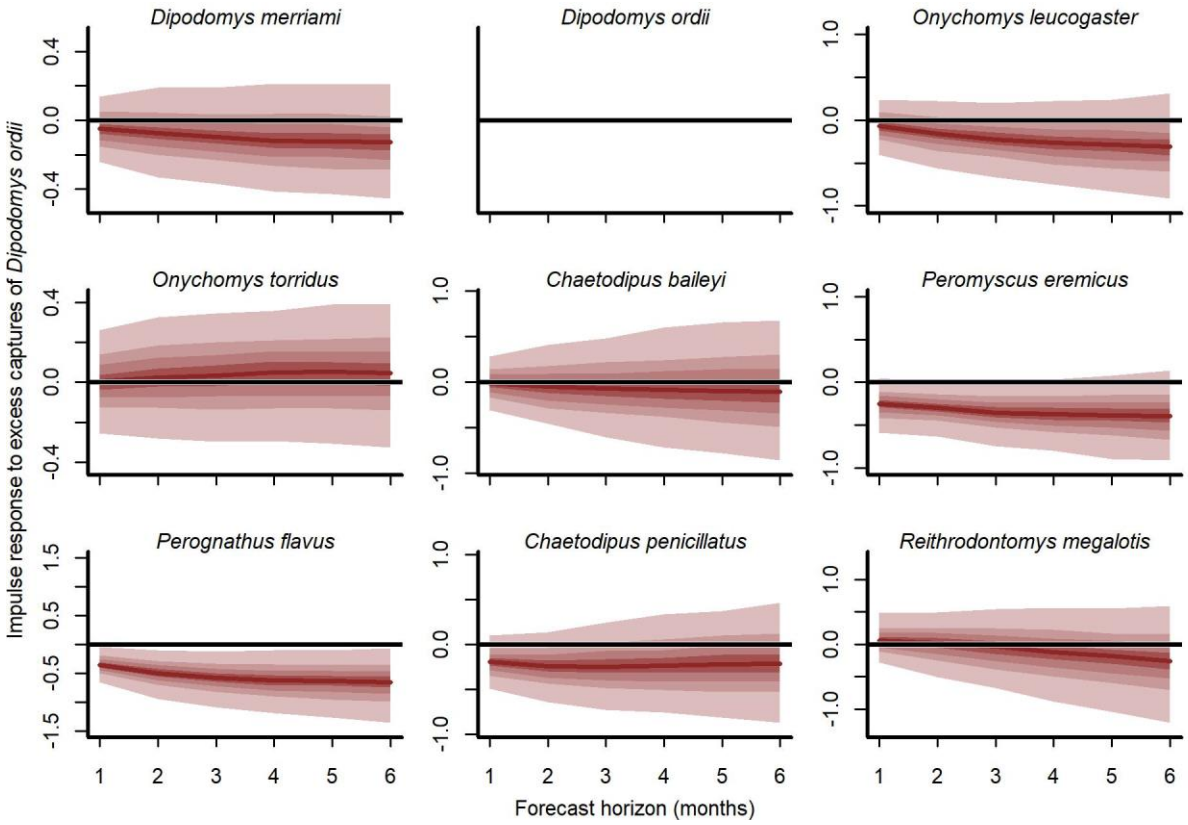


Figure 4: Expected responses to a pulse in captures of Ord's kangaroo rat (*Dipodomys ordii*). Ribbon plots show how mean captures ($\boldsymbol{\mu}$, on the log scale) are expected to change over the next six months if three additional *D. ordii* individuals are captured. Ribbon shading shows posterior empirical quantiles (90th, 60th, 40th and 20th). Dark red lines show posterior medians.

Forecast uncertainties were dominated by uncertainty in the trend component, as opposed to uncertainty in the GAM component (Supplementary Figure S8). This motivated us to ask how trend uncertainty for one species was related to captures of other species. Using a variance decomposition, we computed the relative contributions of pulses from all other species to trend variability for a focal species. Not surprisingly, we found a range of patterns (Supplementary Figure S9). Some species were more tightly related to their own lagged dynamics than to other species. Members of this group included the two grasshopper mouses (*O. leucogaster* and *O. torridus*), silky

pocket mouse (*P. flavus*) and Merriam's kangaroo rat (*D. merriami*). For others, dynamics were governed by a broader suite of interspecies dependencies. For example, imagine we wish to forecast of captures for desert pocket mouse (*C. penicillatus*) on month ahead. According to the model, forecasts would be more sensitive to current captures of Northern grasshopper mouse (*O. leucogaster*) and Western harvest mouse (*R. megalotis*) than to current captures of the desert pocket mouse (Supplementary Figure S9).

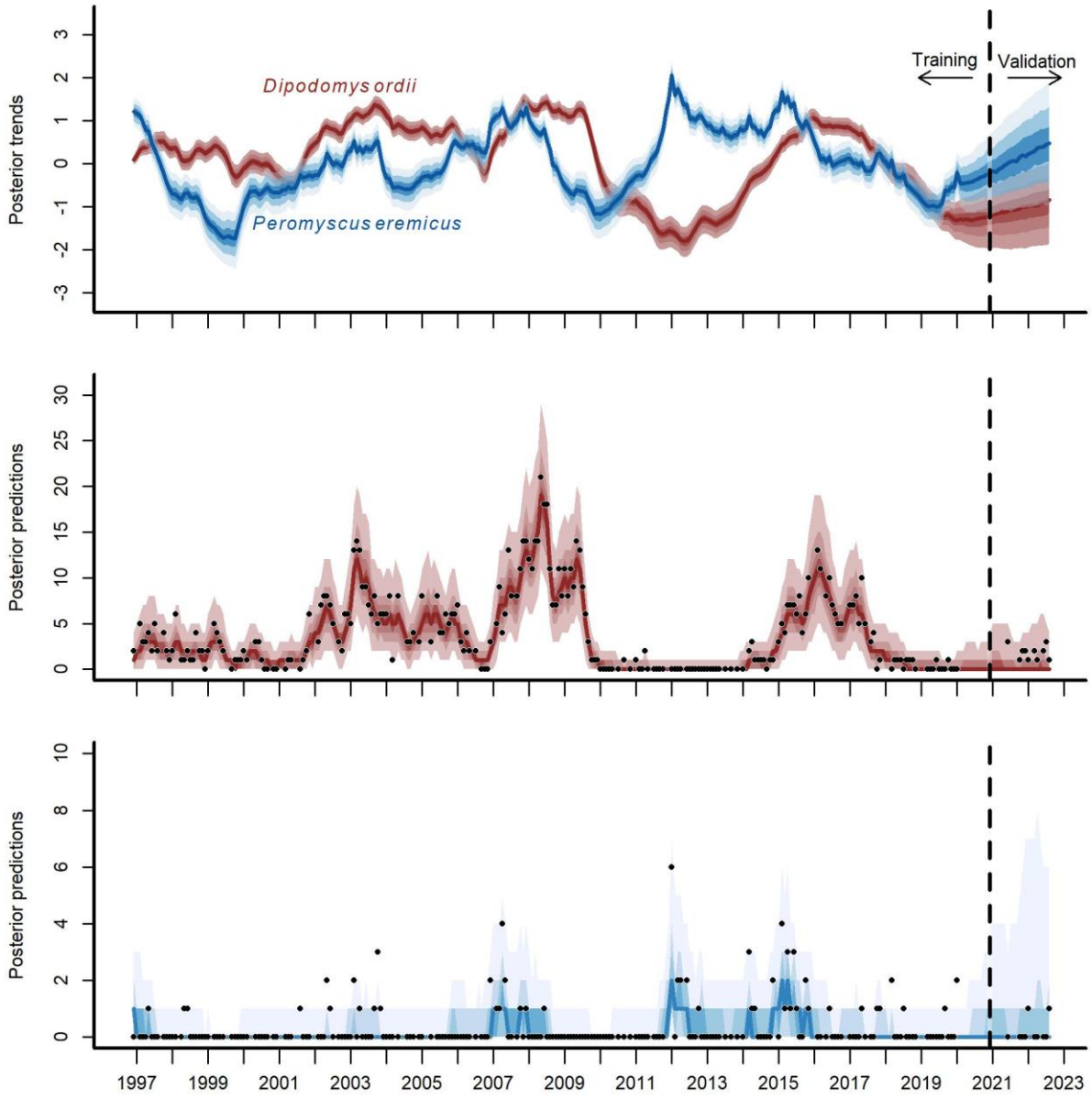
Improved community forecasts

Ten of the 21 validation timepoints had non-missing observations that could be used to compare forecasts. Across these validation points, variogram scores for the three competing models were generally ranked by model complexity. ***GAM-VAR*** forecasts were the most accurate in eight of ten, and always scored better than forecasts from the ***GAM-AR*** (Supplementary Figure S10). Forecasts from the simplest model, ***AR***, were the least accurate in seven of ten validation points. Evaluations did not change when using unscaled forecasts. We visualized posterior hindcast and forecast distributions to better understand how the ***GAM-VAR*** model outperformed the benchmarks. Prediction uncertainties for individual species were well-calibrated, and the VAR process was able to reproduce the observed multi-species temporal dynamics. For example, Ord's kangaroo rat (*D. ordii*) and cactus mouse (*P. eremicus*) had negative cross-dependencies in the ***GAM-VAR***, suggesting their trends should be somewhat structured. Posterior trends and predictions for these species confirm the model learned some of this structure to produce forecasts (Figure 5). The benchmarks produced smoother, less synchronous trends and flatter forecasts (Supplementary Figure S11). Estimates of trend variation were also larger for the benchmarks than the ***GAM-VAR*** for nearly all species (Supplementary Figure S12).

We simulated communities to investigate whether the model could recreate known community transitions. This was done by sampling 200 individual rodents at a given historical

timepoint using multinomial draws. Multinomial sampling weights were chosen based on the model's posterior median expectation, $\mathbf{exp}(\boldsymbol{\mu})$, for each species at each timepoint of interest. Simulated communities accurately reflected community changes that took place during the study period. Notable transitions included the shift to previously 'inferior' competitors from 2000 – 05 following the establishment of Bailey's pocket mouse *C. baileyi*, and the reshuffling that happened following a drought in 2008 – 09 (Supplementary Figure S13).

Posterior estimates of $\boldsymbol{\Phi}$ for most species indicated strong support for overdispersion (Supplementary Figure S14). Inspection of randomized quantile residuals uncovered no obvious evidence of unmodelled temporal or systematic variation (Supplementary Figures S15– S16).



420

421 **Figure 5:** Posterior trend and hindcast/forecast distributions for Ord's kangaroo rat (*Dipodomys*
422 *ordii*; in red) and cactus mouse (*Peromyscus eremicus*; in blue). Trends were scaled to unit variance
423 for comparisons. Points are monthly total observed captures across 10 control plots. Ribbon shading
424 shows posterior empirical quantiles (90th, 60th, 40th and 20th). Dark lines show posterior medians.

425

426 **DISCUSSION**

This paper presents a Bayesian analysis of temporal variation in long-term rodent capture data. Community dynamics in this system were the product of environmental variation and multi-species dependence, as well as other un-modelled factors. Captures for all species increased with higher vegetation greenness and responded nonlinearly to temperature change. But the shapes and magnitudes of these responses differed across species. Biotic structure in trend estimates suggested that capture variation for some species can have cascading community effects, possibly underlying regime transitions. Ignoring one or more of these sources of variation led to less realistic forecasts. Models that describe biological complexity, both through nonlinear covariate functions and multi-species dependence, are useful to ask targeted questions about drivers of change (Ives et al. 2003, Greenville et al. 2016, Pedersen et al. 2019).

Understanding multi-species dependence and forecasting regime transitions

The way a species responds to environmental change depends partly on how this change influences its ability to find resources and reproduce (Heske et al. 1994). But it also arises from effects on the abundances and vital rates of other species (Ives et al. 2003). Our analyses show why models that target both sources of variation should be default when studying community dynamics. The ***GAM-VAR***'s trend variance estimates were smaller than those from the benchmarks because it used more information from the data. It acquired this information from multi-species dependencies, which it used to produce more realistic predictions. What do these dependencies mean? Like other multivariate autoregressive models (Ives et al. 2003, Holmes et al. 2014), the ***GAM-VAR*** is not a biologically plausible model of community dynamics. But as an approximation to more complex models such as a Lotka Volterra system (Volterra 1931), our approach makes it possible to ask ecological questions that would be lost otherwise (Hampton et al. 2013, Holmes et al. 2014, Greenville et al. 2016). Which species have the strongest cascading effects? What changes might we expect if management increases or decreases abundance for target species? How could these effects

relate to regime transitions? An immediate benefit of modeling environmental responses and multispecies dependence jointly is that it is possible to estimate their relative importance in limiting forecast uncertainty. In our study, forecasts were dominated by uncertainty in multi-species trends. But using a vector autoregressive process allowed us to dissect this uncertainty in meaningful ways (Lütkepohl 1990, Ives et al. 2003). Simulated responses to sudden impulses in captures were often delayed. Despite the restriction to a VAR of lag of one month, these responses resulted in cascading changes that lasted up to six months. Variance decompositions pointed to even more complexity. Expected changes for some species were influenced, often nonlinearly, by lagged impulses of multiple other species.

Ecological forecasts often extend over multidecadal scales (Dietze et al. 2018, Clark et al. 2020), but our data and analyses show why this is difficult. The Portal study demonstrates that communities can show unexpected, and sometimes rapid, changes following a variety of disturbances (Brown et al. 1997, Ernest and Brown 2001, Christensen et al. 2018). We restricted our predictions to two years beyond the training data, and they performed well. But these were not real forecasts because we had access to the true environmental measurements for that period. We cannot rely on multi-year forecasts because it is difficult to realistically predict what disturbances or climate changes will happen (Dietze et al. 2018). Forecasts that extend over tens or hundreds of years are also problematic because they undervalue the practice of falsifiable hypothesis-testing (White et al. 2019).

Hierarchical functions of NDVI and minimum temperature

The hierarchical structure of our model assumed that species vary in their responses to NDVI and minimum temperature. But it also allowed the data to influence how much and in what ways this variation occurred. We found positive linear associations between capture rates and a 12-month moving average of NDVI. This was not surprising. The rodents at Portal depend on plants for food

and resources (Ernest et al. 2000, Brown and Ernest 2002), and NDVI measures vegetation greenness in the landscape. But interesting patterns emerged in the variation of these responses. The strongest positive association was shown by Ord's kangaroo rat (*D. ordii*). Field evidence suggests this species consumes and harvests grasses (Kerley et al. 1997), so a strong response to NDVI is a sensible expectation. In contrast, Merriam's kangaroo rat (*D. Merriam*) and desert pocket mouse (*C. penicillatus*) showed some of the weakest associations with NDVI. These species have gained dominance on control plots in recent years as woody shrubs have slowly taken over arid grasslands (Brown et al. 1997). It is difficult to explain why these species would not respond as strongly to greener landscapes. Perhaps they are more capable of surviving in dryer, browner years than other species, which could partially explain their dominance following the 2009 – 10 drought event (Christensen et al. 2018). Or perhaps influxes of other species following periods of increased vegetation greenness result in a landscape that is too competitive for Merriam's kangaroo rat and desert pocket mouse to continue their usual dominance.

Hierarchical distributed lag functions are not common in ecology, but their advantages are numerous. Regularizing species-level responses to change toward a community 'average' response is a powerful technique to improve inferences and predictions (Pedersen et al. 2019, McElreath 2020). In our study, we used hierarchical nonlinear functions to provide useful insights into delayed responses to temperature change for rodents at Portal. Most species showed higher captures when minimum temperatures were low 3 – 4 months prior, suggesting increases begin during mid to late spring when resources such as seeds become available. But others, such as Merriam's kangaroo rat and Southern grasshopper mouse, bucked this trend by increasing during cooler months in autumn and winter. Asynchronous phenology, where species show different reproductive timing, is sometimes expected in competitive communities (Carter and Rudolf 2022). Analysis of individual reproductive status in different biotic contexts suggests some species shift their reproductive timing in the presence of strong competitors in the Portal system (Dumandan et al. 2022). Do these competitive forces play a role in seasonal capture variation over the long-term? Comparing

503 minimum temperature responses on control vs experimental plots would be one interesting way to
504 begin tackling this question.

505 We cannot interpret our environmental response estimates as directly causal, for several
506 reasons. First, we know NDVI is not a perfect measure of changes in seed production. Second, it is
507 likely that changes to NDVI and minimum temperature are both related to other unmeasured
508 environmental drivers that may also influence rodent abundance. Major storms, the El Niño
509 Southern Oscillation and other factors that influence moisture levels can all influence temperature
510 and vegetation change (Sun and Kafatos 2007). These other drivers could act as unmeasured
511 confounds, biasing estimates in a causal inference framework (McElreath 2020).

512

513 **Future directions**

514 We do not know the precise mechanisms that explain our estimated multi-species dependencies.
515 This is a drawback of the model. But it also provides a valuable opportunity to develop hypotheses
516 about the drivers of community change in semi-arid systems. There are several reasons why changes
517 in abundance for one species can result in delayed changes for competitors (Heske et al. 1994,
518 Hampton et al. 2013). Most rodents in the system use seed caches and may be temporarily buffered
519 against changes in competitor abundance (Brown and Munger 1985). Reproduction is also episodic
520 for most species, so we cannot expect an immediate increase in abundance following a decline in
521 competition or predation (Brown and Ernest 2002). A productive avenue for future research could
522 be to gather more detailed environmental measurements to identify the true proximate vegetation-
523 related drivers of population dynamics in this system. Better understanding of these mechanisms
524 could also improve predictions in our modelling framework. Our model produced conditional
525 forecasts, where expected trends for some species influence expected trends for others. This means
526 we could realistically expect improved predictions for the entire community if we can learn more
527 about dynamics for only a few species.

Other work could target model development. For example, we could estimate 9x9 matrices of autoregressive coefficients thanks to Stan's superior Hamiltonian Monte Carlo samplers. But extending our approach to larger species assemblages could be difficult. Dimension reduction using factor models is one possible solution (Warton et al. 2015, Ovaskainen et al. 2017). There is also a need for broader comparison of models to understand whether features of the data can guide model development or informed forecast combinations (Clark et al. 2022, Powell-Romero et al. 2023). This may be particularly useful in situations where prediction accuracy is the primary goal. Developing models that can fuse the decades-worth of valuable pre-existing knowledge that has resulted from the Portal experiments should also be a key focus (Mikkola et al. 2021).

Conclusions

Approaching the challenges of understanding and predicting ecosystem change requires models that enforce realistic biotic structure in near-term ecological forecasts (Hampton et al. 2013, Holmes et al. 2014). Dynamic GAMs provide one possible solution. We hope that the ability to estimate multi-species dependence and species-level variation in nonlinear environmental responses will inspire new questions about the factors that govern ecological community dynamics.

ACKNOWLEDGEMENTS

We thank the many volunteers for their help during fieldwork to generate primary Portal data. This study was supported by an ARC DECRA Fellowship to N.J. Clark (DE210101439). The Portal Project has been funded nearly continuously since 1977 by the National Science Foundation, most recently by DEB-1929730 to S. K. M. Ernest and E.P. White. Development of portal software packages is supported by this NSF grant, NSF grant DEB-1622425 to S. K. M. Ernest, and the Gordon and Betty Moore Foundation's Data-Driven Discovery Initiative through Grant GBMF4563 to E. P. White.

552

553 **SUPPORTING INFORMATION CAPTIONS**

554 **Figure S1**: Total rodent captures from the Portal Project for the period December 1996 to August
555 2022. Counts represent total captures for nine species across 10 control plots, sampled monthly.
556 Blanks represent missing values.

557

558 **Figure S2**: Autocorrelation functions of rodent capture time series in the Portal Project. Dashed lines
559 show values beyond which the autocorrelations are considered significantly different from zero.

560

561 **Figure S3**: Histograms of rodent capture time series in the Portal Project. Counts represent total
562 captures across 10 control plots, sampled monthly.

563

564 **Figure S4**: Seasonal and Trend decomposition using Loess smoothing (STL) applied to observed
565 minimum temperature time series for the period December 1996 – August 2022. The top panel
566 shows the raw time series. The middle plot shows the estimated long-term trend (calculated using a
567 Loess regression to the de-seasoned time series). The bottom plot shows the time-varying estimate
568 of seasonality (calculated using a Loess regression that smooths across years).

569

570 **Figure S5**: Top panel: observed Normalized Difference Vegetation Index (NDVI) time series for the
571 period December 1996 – August 2022, with obvious seasonal fluctuations. Bottom panel: a 12-
572 month moving average that represents smooth, gradual changes in NDVI at the study site.

573

Figure S6: Posterior distributions of vector autoregressive coefficients (matrix $\boldsymbol{A}$). Off-diagonals represent cross-dependencies. For example, the entry in $\boldsymbol{A}[\mathbf{1},\mathbf{2}]$ captures the effect of **DO**'s trend at time $t-1$ on the current trend for **DM** (at time t). Diagonals (with grey shading) represent autoregressive coefficients (the effect of a species' trend at time $t-1$on its own trend at time t). Colours indicate the proportion of probability mass at or below zero (in blue) vs above zero (in red). **DO**, *Dipodomys merriami*; **DO**, *Dipodomys ordii*; **OL**, *Onychomys leucogaster*; **OT**, *Onychomys torridus*; **PB**, *Chaetodipus baileyi*; **PE**, *Peromyscus eremicus*; **PF**, *Perognathus flavus*; **PP**, *Chaetodipus penicillatus*; **RM**, *Reithrodontomys megalotis*.

Figure S7: Expected responses to a pulse in captures of the desert pocket mouse (*Chaetodipus penicillatus*). Ribbon plots show how mean captures (μ, on the log scale) are expected to change over the next six months if three additional *C. penicillatus* individuals are captured. Ribbon shading shows posterior empirical quantiles (90th, 60th, 40th and 20th). Dark red lines show posterior medians.

Figure S8: Relative contributions of uncertainty in the latent trend and GAM components of the linear predictor to forecast uncertainty over increasing forecast horizons.

Figure S9: Latent trend variance decompositions for a few species. Each line shows the relative contribution of a sudden pulse in captures at time zero to the focal species' trend variance over a six-month forecast horizon. Black lines show relative contributions of pulses for the focal species on their own trend variance. Other lines show relative contributions of pulses for the remaining species in the community. Interesting relationships are highlighted in colour. Pulses were simulated as an excess of three captures at time zero.

Figure S10: Out of sample forecast performances of competing models. Y-axis shows the log of the variogram score, a proper score that penalizes multivariate forecasts if they do not capture correlations in observed data. Forecasts were evaluated on 24 out of sample time points (years 2021 and 2022). Points show scores. Lines show Loess smoothed trendlines. Missing values were used for timepoints when sampling did not occur. A lower score indicates a better forecast.

Figure S11: Posterior trend estimates from three competing models for Ord's kangaroo rat (*Dipodomys ordii*; in red) and cactus mouse (*Peromyscus eremicus*; in blue). Trends were scaled to unit variance for comparisons. Ribbon shading shows posterior empirical quantiles (90^{th}, 60^{th}, 40^{th} and 20^{th}). Dark lines show posterior medians.

Figure S12: Posterior estimates of trend standard deviations from the three competing models. Estimates are the square root of diagonal parameters from the trend covariance matrix ($\boldsymbol{\Sigma_{VAR}}$) for the ***GAM-VAR*** (black), ***GAM-AR*** (red) and ***AR*** (blue).

Figure S13: Simulated rodent communities. Using the ***GAM-VAR*** model's posterior predictive distribution, we simulated communities of 200 individuals at different timepoints to investigate how well the model captured known community transitions. Colours represent different species.

Figure S14: Posterior estimates of Negative Binomial overdispersion parameters from the ***GAM-VAR*** (black), ***GAM-AR*** (red) and ***AR*** (blue). Smaller values of Φ indicate a larger amount of overdispersion.

Figure S15: Normal quantile-quantile plots of randomized quantile residuals. Ribbon shading shows posterior empirical quantiles (90th, 60th, 40th and 20th). Dark lines show posterior medians.

Figure S16: Autocorrelation functions of randomized quantile residuals. Ribbon shading shows posterior empirical quantiles (90th, 60th, 40th and 20th). Dark red lines show posterior medians. Dashed lines show values beyond which the autocorrelations would be considered significantly different from zero in a Frequentist paradigm.

DATA AVAILABILITY STATEMENT

Data is available for download using the *portalR* family of packages (Christensen et al. 2019b). R code to reproduce analyses is provided in Supplementary materials and will be permanently archived on Zenodo on acceptance of the manuscript.

REFERENCES

Anderson, C. N., C.-h. Hsieh, S. A. Sandin, R. Hewitt, A. Hollowed, J. Beddington, R. M. May, and G. Sugihara. 2008. Why fishing magnifies fluctuations in fish abundance. Nature **452**:835-839.

Betancourt, M. 2017. A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:1701.02434.

Betancourt, M. 2021. Prior Modelling. Retrieved from: https://github.com/betanalpha/knitr_case_studies/tree/master/prior_modeling, commit 56606fa62e35f87bc88cec6892b4a4d3587f7029.

Biggs, R., S. R. Carpenter, and W. A. Brock. 2009. Turning back from the brink: detecting an impending regime shift in time to avert it. Proceedings of the National Academy of Sciences **106**:826-831.

Bledsoe, E. K., and S. M. Ernest. 2019. Temporal changes in species composition affect a ubiquitous species' use of habitat patches. Ecology **100**:e02869.

Brandt, J. S., M. A. Haynes, T. Kuemmerle, D. M. Waller, and V. C. Radeloff. 2013. Regime shift on the roof of the world: Alpine meadows converting to shrublands in the southern Himalayas. Biological Conservation **158**:116-127.

Brook, B. W., E. C. Ellis, M. P. Perring, A. W. Mackay, and L. Blomqvist. 2013. Does the terrestrial biosphere have planetary tipping points? Trends in Ecology & Evolution **28**:396-401.

Brown, J. H. 1998. The desert granivory experiments at portal. Pages 71-95 *in* W. J. Resetarits and J. Bernardo, editors. Experimental Ecology. Oxford University Press, Oxford, UK.

Brown, J. H., and S. M. Ernest. 2002. Rain and rodents: complex dynamics of desert consumers: although water is the primary limiting resource in desert ecosystems, the relationship between rodent population dynamics and precipitation is complex and nonlinear. BioScience **52**:979-987.

Brown, J. H., and E. J. Heske. 1990. Temporal changes in a Chihuahuan Desert rodent community. Oikos:290-302.

Brown, J. H., and J. C. Munger. 1985. Experimental manipulation of a desert rodent community: food addition and species removal. Ecology **66**:1545-1563.

Brown, J. H., T. J. Valone, and C. G. Curtin. 1997. Reorganization of an arid ecosystem in response to recent climate change. Proceedings of the National Academy of Sciences **94**:9729-9733.

Brown, J. H., and Z. Zeng. 1989. Comparative population ecology of eleven species of rodents in the Chihuahuan Desert. Ecology **70**:1507-1525.

Carpenter, B., A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: A probabilistic programming language. Journal of Statistical Software **76**.

669 Carpenter, S. R., J. J. Cole, M. L. Pace, R. Batt, W. A. Brock, T. Cline, J. Coloso, J. R. Hodgson, J. F.
670 Kitchell, and D. A. Seekell. 2011. Early warnings of regime shifts: a whole-ecosystem
671 experiment. Science **332**:1079-1082.
672 Carter, S. K., and V. H. Rudolf. 2022. Exploring conditions that strengthen or weaken the ecological
673 and evolutionary consequences of phenological synchrony. The American Naturalist
674 **200**:E189-E206.
675 Chaparro-Pedraza, P. C., and A. M. de Roos. 2020. Ecological changes with minor effect initiate
676 evolution to delayed regime shifts. Nature Ecology & Evolution **4**:412-418.
677 Christensen, E. M., D. J. Harris, and S. Ernest. 2018. Long-term community change through multiple
678 rapid transitions in a desert rodent community. Ecology **99**:1523-1529.
679 Christensen, E. M., G. L. Simpson, and S. Ernest. 2019a. Established rodent community delays
680 recovery of dominant competitor following experimental disturbance. Proceedings of the
681 Royal Society B: Biological Sciences **286**:20192269.
682 Christensen, E. M., G. M. Yenni, H. Ye, J. L. Simonis, E. K. Bledsoe, R. Diaz, S. D. Taylor, E. P. White,
683 and S. Ernest. 2019b. portalr: an R package for summarizing and using the Portal Project
684 Data. Journal of Open Source Software **4**:1098.
685 Cihlar, J., L. S.-. Laurent, and J. Dyer. 1991. Relation between the normalized difference vegetation
686 index and ecological variables. Remote Sensing of Environment **35**:279-298.
687 Clark, J. S., S. R. Carpenter, M. Barber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, M. Pascual, R.
688 Pielke, and W. Pizer. 2001. Ecological forecasts: an emerging imperative. Science **293**:657-
689 660.
690 Clark, N. J., J. T. Kerry, and C. I. Fraser. 2020. Rapid winter warming could disrupt coastal marine fish
691 community structure. Nature Climate Change:DOI: 10.1038/s41558-41020-40838-41555.
692 Clark, N. J., T. Proboste, G. Weerasinghe, and R. J. Soares Magalhães. 2022. Near-term forecasting of
693 companion animal tick paralysis incidence: An iterative ensemble model. PLoS
694 Computational Biology **18**:e1009874.

695 Clark, N. J., and K. Wells. 2022. Dynamic generalised additive models (DGAMs) for forecasting
696 discrete ecological time series. Methods in Ecology and Evolution **00**:1-14.
697 Cosentino, B. J., R. L. Schooley, B. T. Bestelmeyer, J. F. Kelly, and J. M. Coffman. 2014. Constraints
698 and time lags for recovery of a keystone species (Dipodomys spectabilis) after landscape
699 restoration. Landscape ecology **29**:665-675.
700 Diaz, R. M., and S. M. Ernest. 2022. Maintenance of community function through compensation
701 breaks down over time in a desert rodent community. Ecology **103**.
702 Dickman, C., P. Mahon, P. Masters, and D. Gibson. 1999. Long-term dynamics of rodent populations
703 in arid Australia: the influence of rainfall. Wildlife Research **26**:389-403.
704 Dietze, M. C., A. Fox, L. M. Beck-Johnson, J. L. Betancourt, M. B. Hooten, C. S. Jarnevich, T. H. Keitt,
705 M. A. Kenney, C. M. Laney, and L. G. Larsen. 2018. Iterative near-term ecological forecasting:
706 Needs, opportunities, and challenges. Proceedings of the National Academy of Sciences
707 **115**:1424-1432.
708 Dumandan, P. K. T., G. M. Yenni, and M. Ernest. 2022. Shifts in competitive structures can drive
709 variation in species phenology. bioRxiv:2022.2012. 2013.520270.
710 Ernest, S., and J. H. Brown. 2001. Delayed compensation for missing keystone species by
711 colonization. Science **292**:101-104.
712 Ernest, S., G. M. Yenni, G. Allington, E. K. Bledsoe, E. M. Christensen, R. M. Diaz, K. Geluso, J. R.
713 Goheen, Q. Guo, and E. Heske. 2020. The Portal Project: a long-term study of a Chihuahuan
714 desert ecosystem. bioRxiv:332783.
715 Ernest, S. M., J. H. Brown, and R. R. Parmenter. 2000. Rodents, plants, and precipitation: spatial and
716 temporal dynamics of consumers and resources. Oikos **88**:470-482.
717 Foley, J. A., M. T. Coe, M. Scheffer, and G. Wang. 2003. Regime shifts in the Sahara and Sahel:
718 interactions between ecological and climatic systems in Northern Africa. Ecosystems **6**:524-
719 532.

Fukami, T., and D. A. Wardle. 2005. Long-term ecological dynamics: reciprocal insights from natural and anthropogenic gradients. Proceedings of the Royal Society B: Biological Sciences **272**:2105-2115.

Gabry, J., and R. Češnovar. 2021. Cmdstanr: R interface to 'CmdStan'. https://mc-stan.org/cmdstanr.

Gelman, A. 2006. Multilevel (hierarchical) modeling: what it can and cannot do. Technometrics **48**:432-435.

Green, J. L., A. Hastings, P. Arzberger, F. J. Ayala, K. L. Cottingham, K. Cuddington, F. Davis, J. A. Dunne, M.-J. Fortin, and L. Gerber. 2005. Complexity in ecology and conservation: mathematical, statistical, and computational challenges. BioScience **55**:501-510.

Greenville, A. C., G. M. Wardle, V. Nguyen, and C. R. Dickman. 2016. Population dynamics of desert mammals: similarities and contrasts within a multispecies assemblage. Ecosphere **7**:e01343.

Hampton, S. E., E. E. Holmes, L. P. Scheef, M. D. Scheuerell, S. L. Katz, D. E. Pendleton, and E. J. Ward. 2013. Quantifying effects of abiotic and biotic drivers on community dynamics with multivariate autoregressive (MAR) models. Ecology **94**:2663-2669.

Heske, E. J., J. H. Brown, and S. Mistry. 1994. Long-term experimental study of a Chihuahuan Desert rodent community: 13 years of competition. Ecology **75**:438-445.

Holmes, E., E. Ward, and M. Scheuerell. 2014. Analysis of multivariate time-series using the MARSS package. NOAA Fisheries, Northwest Fisheries Science Center **2725**:98112.

Hughes, T. P., C. Linares, V. Dakos, I. A. Van De Leemput, and E. H. Van Nes. 2013. Living dangerously on borrowed time during slow, unrecognized regime shifts. Trends in Ecology & Evolution **28**:149-155.

Ives, A. R., B. Dennis, K. L. Cottingham, and S. R. Carpenter. 2003. Estimating community stability and ecological interactions from time-series data. Ecological Monographs **73**:301-330.

Kenagy, G., and G. A. Bartholomew. 1981. Effects of day length, temperature, and green food on testicular development in a desert pocket mouse Perognathus formosus. Physiological Zoology **54**:62-73.

Kerley, G. I., W. G. Whitford, and F. R. Kay. 1997. Mechanisms for the keystone status of kangaroo rats: graminivory rather than granivory? Oecologia **111**:422-428.

Lasky, J. R., M. B. Hooten, and P. B. Adler. 2020. What processes must we understand to forecast regional-scale population dynamics? Proceedings of the Royal Society B: Biological Sciences **287**:20202219.

Levin, S., T. Xepapadeas, A.-S. Crépin, J. Norberg, A. De Zeeuw, C. Folke, T. Hughes, K. Arrow, S. Barrett, and G. Daily. 2013. Social-ecological systems as complex adaptive systems: modeling and policy implications. Environment and Development Economics **18**:111-132.

Lewis, A. S. L., W. M. Woelmer, H. L. Wander, D. W. Howard, J. W. Smith, R. P. McClure, M. E. Lofton, N. W. Hammond, R. S. Corrigan, R. Q. Thomas, and C. C. Carey. 2022. Increased adoption of best practices in ecological forecasting enables comparisons of forecastability. Ecological Applications **32**:e2500.

Luis, A. D., R. J. Douglass, J. N. Mills, and O. N. Bjørnstad. 2010. The effect of seasonality, density and climate on the population dynamics of Montana deer mice, important reservoir hosts for Sin Nombre hantavirus. Journal of Animal Ecology **79**:462-470.

Lütkepohl, H. 1990. Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models. The review of economics and statistics:116-125.

May, R. M. 1977. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. Nature **269**:471-477.

McElreath, R. 2020. Statistical rethinking: A Bayesian course with examples in R and Stan. 2 edition. Chapman and Hall/CRC.

McElreath, R., and J. Koster. 2014. Using multilevel models to estimate variation in foraging returns. Human Nature **25**:100-120.

Mikkola, P., O. A. Martin, S. Chandramouli, M. Hartmann, O. A. Pla, O. Thomas, H. Pesonen, J. Corander, A. Vehtari, and S. Kaski. 2021. Prior knowledge elicitation: The past, present, and future. arXiv preprint arXiv:2112.01380.

Miller, D. L. 2019. Bayesian views of generalized additive modelling. arXiv preprint arXiv:1902.01330.

Ovaskainen, O., G. Tikhonov, D. Dunson, V. Grøtan, S. Engen, B.-E. Sæther, and N. Abrego. 2017. How are species interactions structured in species-rich communities? A new method for analysing time-series data. Proceedings of the Royal Society B: Biological Sciences **284**:20170768.

Pedersen, E. J., D. L. Miller, G. L. Simpson, and N. Ross. 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ **7**:e6876.

Pettorelli, N., S. Ryan, T. Mueller, N. Bunnefeld, B. Jędrzejewska, M. Lima, and K. Kausrud. 2011. The Normalized Difference Vegetation Index (NDVI): unforeseen successes in animal ecology. Climate research **46**:15-27.

Powell-Romero, F., N. M. Fountain-Jones, A. Norberg, and N. J. Clark. 2023. Improving the predictability and interpretability of co-occurrence modelling through feature-based joint species distribution ensembles. Methods in Ecology and Evolution **14**:146-164.

R Core Team. 2020. R: A language and environment for statistical computing. R Development Core Team, Vienna, Austria.

Roughgarden, J., and F. Smith. 1996. Why fisheries collapse and what to do about it. Proceedings of the National Academy of Sciences **93**:5078-5083.

Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara. 2009. Early-warning signals for critical transitions. Nature **461**:53-59.

Scheuerer, M., and T. M. Hamill. 2015. Variogram-based proper scoring rules for probabilistic forecasts of multivariate quantities. Monthly Weather Review **143**:1321-1334.

Simonis, J. L., E. P. White, and S. K. M. Ernest. 2021. Evaluating probabilistic ecological forecasts. Ecology **102**:e03431.

796 Simonis, J. L., G. M. Yenni, E. K. Bledsoe, E. M. Christensen, H. Senyondo, S. D. Taylor, H. Ye, E. P.
797 White, and S. M. Ernest. 2022. portalcasting: Supporting automated forecasting of rodent
798 populations. Journal of Open Source Software **7**:3220.
799 Simpson, D., H. Rue, A. Riebler, T. G. Martins, and S. H. Sørbye. 2017. Penalising model component
800 complexity: A principled, practical approach to constructing priors. Statistical Science **32**:1-
801 28.
802 Stan Development Team. 2022. Stan Modeling Language Users Guide and Reference Manual,
803 Version 2.26.1. https://mc-stan.org.
804 Sun, D., and M. Kafatos. 2007. Note on the NDVI-LST relationship and the use of temperature-
805 related drought indices over North America. Geophysical Research Letters **34**.
806 Supp, S. R., D. N. Koons, and S. M. Ernest. 2015. Using life history trade-offs to understand core-
807 transient structuring of a small mammal community. Ecosphere **6**:1-15.
808 Vehtari, A., A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner. 2021. Rank-normalization,
809 folding, and localization: an improved R for assessing convergence of MCMC (with
810 discussion). Bayesian Analysis **16**:667-718.
811 Volterra, V. 1931. Variations and fluctuations of the number of individuals in animal species living
812 together. McGraw-Hill.
813 Warton, D. I., F. G. Blanchet, R. B. O'Hara, O. Ovaskainen, S. Taskinen, S. C. Walker, and F. K. Hui.
814 2015. So many variables: joint modeling in community ecology. Trends in Ecology &
815 Evolution **30**:766-779.
816 Wells, K., R. B. O'Hara, B. D. Cooke, G. J. Mutze, T. A. A. Prowse, and D. A. Fordham. 2016.
817 Environmental effects and individual body condition drive seasonal fecundity of rabbits:
818 identifying acute and lagged processes. Oecologia **181**:853-864.
819 White, E. P., G. M. Yenni, S. D. Taylor, E. M. Christensen, E. K. Bledsoe, J. L. Simonis, and S. Ernest.
820 2019. Developing an automated iterative near-term forecasting system for an ecological
821 study. Methods in Ecology and Evolution **10**:332-344.

822 Wood, S. 2017. Generalized additive models: an introduction with R. Second edition. CRC Press, Boco
823 Raton.

824