

SI Appendix: Modeling and computation

The emergent interactions that govern biodiversity change

James S. Clark^{1,2}, C. Lane Scher¹, and Margaret Swift¹

¹*Nicholas School of the Environment, Duke University, Durham NC 27708, USA*

²*Department of Statistical Science, Duke University, Durham NC 27708, USA*

Contents

S1	MODELING BACKGROUND	1
S2	THEORETICAL BACKGROUND	2
A	TECHNICAL BACKGROUND	3
B	A TWO-SPECIES EXAMPLE	4
S3	DATA SIMULATION	4
S4	CONTRIBUTIONS TO DYNAMICS	5
S5	INTERACTIONS AND INDIRECT EFFECTS	6
S6	PRIOR KNOWLEDGE	7
S7	POSTERIOR SIMULATION	7
A	<i>Sampling sparse coefficient matrices</i>	7
B	<i>Sampling latent states</i>	8
S8	APPLICATIONS TO FOOD WEB DATA	8
A	<i>Wisconsin Experimental Lakes</i>	8
B	<i>Breeding bird survey</i>	10

S1. MODELING BACKGROUND

gjamTime is built from the data model of Generalized Joint Attribute Modeling (GJAM) (1), embedding it within a dynamic framework that includes environment-species interactions (ESI) and resolves several issues posed by multivariate, time-series, species-abundance data. The challenges of fitting Lotka-Volterra (LV) models to data have limited direct tests of theory and their application (2, 3). Fitting static versions of LV to data omits the serial dependence in observations and the errors in process and observations (4, 5). This dependence between observations is needed to estimate uncertainty, which is especially important where data are noisy. *Process error* allows for model miss-specification, the fact that LV is a crude approximation at best. Due to the time-series nature of dynamics, process error accumulates over time. *Observation error* allows for the difference between what is observed and the 'true' population size. Populations in Figure 2 of

the main text number in the millions of individuals, but may be represented in data by counts in single digits. Due to combined errors, the fraction of total variance explained by a model like LV could be small. [The LV model is summarized in section S2.] If a model explains, say, 10% of the variance in dynamic data, the signal may be too weak to estimate parameters, make useful predictions, or test hypotheses. The stochastic treatment of the underlying process that is possible in models that include both process and observation error allows for the conditional-independence assumption of serial data. The conditionally independent data become marginally dependent when considered jointly with the latent states. This hierarchical treatment allows for valid credible intervals on parameters and valid predictive intervals on unobserved states (4, 6).

There are many curve-fitting algorithms applied to time-series data that we do not attempt to review here. The observation-process error combination that is missing from many of the more algorithmic approaches (7–9) is needed to allow for serial dependence in observations. Algorithmic approaches often predict point values as well as full probability models, but lack a probability basis for credible and predictive intervals. In addition, the process model used here is motivated by the need for a bio-physical interpretation of parameter estimates.

Especially relevant are the recent advances in the AR(1) models introduced by (2, 10, 11). A growth-rate parameter ρ in a population model that links population size $w_{s,t}$ of species s at time t to $w_{s,t+1}$ is an AR(1) coefficient. Recent AR(1) models applied to community data further allow that the growth rate of species s depends on the abundance of other species in the vector \mathbf{w}_t . For a stationary process, the AR(1) model can be viewed as an approximation to the non-linear, generalized LV model in the neighborhood of an equilibrium \mathbf{w}^* (2). For communities that exhibit non-stationary dynamics, and abundances commonly move far from equilibrium values, the linear AR(1) and the non-linear LV models are expected to diverge. The feedback of population abundance $w(t)$ on population growth rate dw/dt , through a quadratic term containing w^2 is omitted from AR models and poses computational challenges.

The challenge of many zeros in species abundance data that motivated GJAM is important for time-series data. A preponderance of zeros can restrict analysis to data where

zeros do not occur. It can motivate a non-linear link function that complicates interpretation. For example, where data are modeled on a log scale, it is not possible to state the effect of a treatment, a covariate, or other species s' on the growth rate of s without qualifying it by saying how abundant the species are. There is no interpretation of 'main effects'; on the log scale, all effects are multiplicative (they are interactions). Where scale differences are side-stepped by treating abundances as presence/absence, most of the information in data is lost, without addressing the fact that the zeros and ones still depend on the differing (and usually non-comparable) effort applied to species observed in different ways.

gjamTime moves from the LV model of eqn 1 to a form that is suited to model fitting in eq. (S2.5). It contains elements of several models in the literature. Using the first term of eqn 1 alone, it is (static) GJAM (1). Omitting the first term in eqn 1 follows the basic LV assumptions. Examples of dynamic models for inference on individual species include (5, 6, 12). In the absence of the GJAM data model (section A), the first two terms are a joint autoregressive AR(1) model, but differing from (2) and (11) in that gjamTime models dynamics on the observation scale. The ordinal model of (10) can be close to the observation scale, depending in part on the ordinal scale that is used. The (static) GJAM model adopts elements of censoring in (13, 14), as discussed in (1).

Allowing for observation error in gjamTime introduces a latent process that absorbs serial dependence that is not limited to AR(1). As mentioned above, this latent process permits the assumption of conditional independence in observations (4, 15); there will be posterior covariance in latent states that could include a rich lag-covariance structure, as would be expected from populations having age or size structure and interacting with unobserved species and environmental variables. This covariance is available from the posterior distribution.

gjamTime additionally includes density-dependence, i.e., the effect of species on one another that depends on their combined abundances (the third term of eqn 1). Even with hierarchical Bayesian analysis this term comes with computational difficulties that are addressed in section S2.

S2. THEORETICAL BACKGROUND

The generalized LV model describes how the abundances of species affect growth rates of one another. The dynamics of a species s are described by two terms, i) an autoregressive density-independent growth rate and ii) feedback from S other species w_1, \dots, w_S . The strength of these interactions is determined by coefficients in a Jacobian matrix α ,

$$\frac{dw_s(t)}{dt} = w_s(t) \left(\rho_s + \sum_{s'=1}^S \alpha_{ss'} w_{s'}(t) \right)$$

or, written as a system of S equations,

$$\frac{d\mathbf{w}(t)}{dt} = \text{diag}(\mathbf{w}(t)) [\boldsymbol{\rho} + \boldsymbol{\alpha}\mathbf{w}(t)] \quad [\text{S2.1}]$$

where $\mathbf{w}(t)$ is a length- S vector of species abundances, and $\text{diag}(\mathbf{w}(t))$ is a $S \times S$ matrix of zeros with elements of $\mathbf{w}(t)$

arranged along the diagonal. The *density-independent (DI) growth coefficients* in the first term are held in a vector $\boldsymbol{\rho} = (\rho_1, \dots, \rho_S)$. DI growth is positive if $\rho_s > 0$. The *species interaction coefficients* $\alpha_{ss'}$ in the $S \times S$ matrix $\boldsymbol{\alpha}$ describe the effect of the abundance of species s' on the growth rate of species s (16–19). For competition, $\alpha_{ss'}$ is negative. For predator s and prey s' , $\alpha_{ss'} > 0$ (a resource that stimulates growth), and $\alpha_{s's} < 0$ (predation loss). We expect asymmetry, $\alpha_{ss'} \neq \alpha_{s's}$, because two species cannot have precisely the same effects on one another. The *carrying capacity* for the species in the absence of other species depends only on its own parameter values

$$w_s^c = -\frac{\rho_s}{\alpha_{ss}} \quad [\text{S2.2}]$$

The *equilibrium community* has abundances that each depend on one another through the species interaction matrix $\boldsymbol{\alpha}$,

$$\mathbf{w}^* = -(\boldsymbol{\alpha}\boldsymbol{\alpha}')^{-1}\boldsymbol{\alpha}'\boldsymbol{\rho} \quad [\text{S2.3}]$$

A necessary (but not sufficient) condition for coexistence of competing species (i.e., all $w_s^* > 0$) is an interaction matrix $\boldsymbol{\alpha}$ that has *diagonal dominance*, i.e., there is a tendency for species to disproportionately limit their own growth rates ($\alpha_{ss} \ll 0$). More generally, there is *local stability* if all eigenvalues of the Jacobian matrix $\boldsymbol{\alpha}$ have negative real parts (16, 20). Dynamics can be unpredictable, because the relationship between two species s and s' at time t affects the growth rate of every other species in the future, through all coefficients in $\boldsymbol{\alpha}$. We refer to these effects as the *conditional* (or 'direct') effects in $\boldsymbol{\alpha}$ and *marginal* (or 'direct plus indirect') effects, respectively.

The interaction coefficients of eq. (S2.1) are held in a $S \times S$ matrix $\boldsymbol{\alpha}$,

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1S} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{S1} & \alpha_{S2} & \dots & \alpha_{SS} \end{pmatrix}$$

The structure of $\boldsymbol{\alpha}$ is the subject of an extensive literature on trophic interactions, food webs, and species coexistence. Trophic (food-web) theory is often organized around this *community matrix* of interactions, which summarizes the effects of each species on growth rates of others (21). For two species, the community matrix is

$$\boldsymbol{\alpha}_= = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$$

In food web models, interactions are often summarized by the signs of interactions. Inter- and intraspecific competition can be represented as

$$\boldsymbol{\alpha}_\pm = \begin{pmatrix} - & - \\ - & - \end{pmatrix}$$

and predation as

$$\alpha_{\pm} = \begin{pmatrix} - & + \\ - & - \end{pmatrix}$$

Amensalism (one-way) competition is

$$\alpha_{\pm} = \begin{pmatrix} - & 0 \\ - & - \end{pmatrix}$$

Mutualism is

$$\alpha_{\pm} = \begin{pmatrix} - & + \\ + & - \end{pmatrix}$$

Beyond the signs of the coefficients there is substantial literature on their relative magnitudes, sometimes termed *interaction strength*. For competitors ($\alpha_{ss'} \leq 0$) the question may focus on whether or not diagonal elements are sufficiently negative to insure coexistence. Do competing sessile organisms recognize the identities of their neighbors (all $\alpha_{ss'}$ unique), only some, or none (all $\alpha_{ss'}$ equivalent)? Do competitors affect one another directly ($\alpha_{ss'} \neq 0$) or by way of other species with which they interact? These questions, and many related ones, depend on a way to estimate their magnitudes, with full uncertainty.

Table S2.1. Variables and parameters in gjamTime.

Name ^a	Dimensions ^b	Definition ^c
Variables		
\mathbf{u}_t	$U \times 1$	$\tilde{\mathbf{w}}_{t-1} \otimes \tilde{\mathbf{w}}_t$ with all combinations $\tilde{w}_{s,t} \cdot \tilde{w}_{s',t}$
\mathbf{v}_t	$V \times 1$	$\tilde{\mathbf{w}}_{t-1} \otimes \mathbf{x}_t$ with all combinations $\tilde{w}_{s,t} \cdot x_{q,t}$
\mathbf{w}_t	$S \times 1$	(latent) species abundance per effort for $w_{1,t}, \dots, w_{S,t}$
$\tilde{\mathbf{w}}_t$	$S \times 1$	censored abundance per effort with elements $\max(0, w_{s,t})$
\mathbf{x}_t	$Q \times 1$	design vector for $x_{1,t}, \dots, x_{Q,t}$
\mathbf{y}_t	$S \times 1$	observed abundance $y_{1,t}, \dots, y_{S,t}$
$\boldsymbol{\epsilon}_t$	$S \times 1$	error vector
Lotka Volterra parameters		
$\boldsymbol{\alpha}$	$S \times S$	spp interactions $\alpha_{s,s'}$ for effects of s' on growth of s
$\boldsymbol{\rho}$	$Q \times S$	DI growth with elements $\rho_{q,s}$
$\boldsymbol{\beta}$	$Q \times S$	movement coefficients $\beta_{q,s}$, multiplies \mathbf{x}_t
Fitted model		
\mathbf{P}	$V \times S$	$\boldsymbol{\rho}$ reorganized as in eq. (S2.9) multiplies \mathbf{v}_t
\mathbf{A}	$U \times S$	$\boldsymbol{\alpha}$ reorganized as in eq. (S2.10) multiplies \mathbf{u}_t
$\boldsymbol{\Sigma}$	$S \times S$	residual covariance

^a The latent variable $w_{s,t}$ can be negative when $y_{s,t}$ is zero, but its effect on population growth is non-negative, $\tilde{w}_{s,t} = \max(0, w_{s,t})$; it is a censored version of $w_{s,t}$

^b S - number of species; Q - number of predictors; U - less than the number of species combinations S^2 ; V - less than the number of species-predictor combinations SQ

^c The symbol \otimes is the Kronecker product.

A. TECHNICAL BACKGROUND. The discrete-time, stochastic version of the continuous LV model (eqn 1) describes change over an interval of duration dt ,

$$\begin{aligned} \mathbf{w}_{t+dt} - \mathbf{w}_t &= dt \cdot (d\mathbf{w}_t + \boldsymbol{\Sigma}^{1/2} \boldsymbol{\epsilon}_t) \\ d\mathbf{w}_t &= \boldsymbol{\beta}' \mathbf{x}_t + \mathbf{P}' \mathbf{v}_t + \mathbf{A}' \mathbf{u}_t \end{aligned} \quad [\text{S2.4}]$$

where $d\mathbf{w}_t$ is a discrete-time version of the differential equation in eqn 1, $\boldsymbol{\epsilon}_t$ is a random vector, and $\boldsymbol{\Sigma}^{1/2}$ is the square-root matrix for the $S \times S$ process error covariance. In addition to time t there will usually be a location i subscript that is omitted here to reduce clutter. The first term includes $S \times Q$ matrix $\boldsymbol{\beta}$, as defined previously. The length- Q vector of predictors $\mathbf{x}_t = (1, x_{2,t}, \dots, x_{Q,t})'$ includes an intercept. The second term includes environment-species interactions $w_{s,t} \cdot x_{q,t}$ in length- V vector \mathbf{v}_t , where V is a subset of all possible combinations of species abundances in \mathbf{w}_t and environmental variables in \mathbf{x}_t . \mathbf{P} is a sparse matrix that reorganizes $\boldsymbol{\rho}$ to optimize posterior simulation (see below). The third term holds species interactions $w_{s,t} \cdot w_{s',t}$ in length- U vector \mathbf{u}_t , where U is less than the number of species combinations. \mathbf{A} holds the interaction coefficients from $\boldsymbol{\alpha}$ in a matrix that is also sparse and reorganized to optimize model fitting (section A). Coefficients and variables are defined in table S2.1. For $dt = 1$, eq. (S2.4) becomes

$$\Delta \mathbf{w}_t = \boldsymbol{\beta}' \mathbf{x}_t + \mathbf{P}' \mathbf{v}_t + \mathbf{A}' \mathbf{u}_t + \boldsymbol{\Sigma}^{1/2} \boldsymbol{\epsilon}_t \quad [\text{S2.5}]$$

where $\Delta \mathbf{w}_t = \mathbf{w}_{t+1} - \mathbf{w}_t$ is a growth increment for population abundances of S species.

The model eq. (S2.5) is combined with data as in GJAM, with the important distinction that the model now operates not on \mathbf{w}_i , but rather on its change with time. An observation consists of vectors $\{\mathbf{x}_{i,t}, \mathbf{y}_{i,t}\}$ at location ('site') $i = 1, \dots, n$, on sample occasion $t = 1, \dots, T_i$. The length- S response vector $\mathbf{y}_{i,t} = (y_{i,t,1} \dots y_{i,t,S})'$ can include species observed as discrete counts or continuous abundances. In the current version of gjamTime they are discretized to allow for errors in observation, which are determined by the width of discrete bins and sampling effort. Count $\mathbf{y}_{i,t}$ is a discrete version of a latent (continuous) variable $\mathbf{w}_{i,t}$.

$$\begin{aligned} \Delta \mathbf{w}_{i,t} | \mathbf{x}_{i,t}, \mathbf{w}_{i,t}, \mathbf{y}_{i,t+1} &\sim MVN(d\mathbf{w}_{i,t}, \boldsymbol{\Sigma}) \times \prod_{s=1}^S \mathcal{I}_{s,i,t} \\ \mathcal{I}_{s,i,t} &= \prod_{k \in \mathcal{C}} I_{s,i,t,k}^{1(y_{s,i,t+1}=k)} (1 - I_{s,i,t,k})^{1(y_{s,i,t+1} \neq k)} \end{aligned} \quad [\text{S2.6}]$$

where the indicator function $1(\cdot)$ is equal to 1 when its argument is true and zero otherwise (1, 22). The indicator

$$I_{s,i,t,k} = 1(p_{s,i,t+1,k} < w_{s,i,t+1} < p_{s,i,t+1,k+1})$$

means that the abundance changes over the interval $(t, t+1)$ to a value $w_{s,i,t+1}$ that lies within the interval of the partition that is assigned to the observed value $y_{s,i,t+1}$. The partition depends on effort,

$$\mathbf{p}_{s,i,t,k} = \left(\frac{k-1/2}{E_{s,i,t}}, \frac{k+1/2}{E_{s,i,t}} \right) \quad [\text{S2.7}]$$

where $E_{s,i,t}$ is the effort applied to observe species s at i, t . For plants, effort can be plot area. For point counts, it can be distance traveled, observation time, or their product. Large effort makes intervals narrow, which, in turn, reduce observation variance, thereby increasing the weight of the observation in the model fit. To reduce clutter, we hereafter omit the location i subscript. The model graph is shown in [fig. S2.1](#).

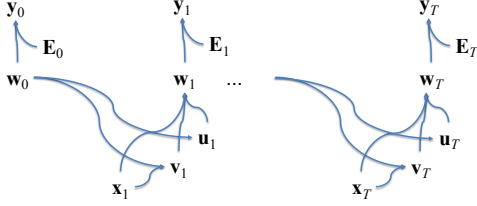


Fig. S2.1. Model graph with subscripts for time (location is omitted). Observed responses are y_1, \dots, y_T . Observation effort is given by $E_{t,s}$. The first vector y_0 and missing values are assigned prior values, the weight of which is determined by the effort that is assigned to them, through $E_{s,t}$.

B. A TWO-SPECIES EXAMPLE. For simplicity we use a two-species example to discuss elements of the model, where abundances occupy the vector $\mathbf{w}_t = (w_{1,t}, w_{2,t})$. Corresponding to the length- Q vector of environmental variables ('predictors') for movement \mathbf{x}_t , there is a $Q \times S$ coefficient matrix

$$\beta = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \vdots & \vdots \\ \beta_{Q1} & \beta_{Q2} \end{pmatrix}$$

Again, the term $\beta' \mathbf{x}_t$ in [eq. \(S2.5\)](#) is independent of population densities.

To optimize posterior simulation we reorganize DI matrix ρ and DD matrix α . Environment is expected to enter not only through its effects on movement, but also on population growth rate. For DI effects, there is a matrix of coefficients for environment-species interactions (ESI) that need not include the same predictors in \mathbf{x} (see [section S8](#)). To avoid introducing more variables we use \mathbf{x} to represent environmental variables in both of the first two terms of [eq. \(S2.5\)](#). The two-species version of the matrix for effects on DI growth has the $Q \times 2$ matrix ρ

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \vdots & \vdots \\ \rho_{Q1} & \rho_{Q2} \end{pmatrix} \quad [\text{S2.8}]$$

For each species s there are parameters $\rho_{1,s}, \dots, \rho_{Q,s}$ for the interactions between abundance and environmental variables $w_{s,t} \cdot x_{q,t}$. There are QS possible interactions between S species and Q predictors in ρ . However, not all Q variables in \mathbf{x}_t may influence population growth of all species S , so

$V \leq QS$. Potential elements are drawn from the vector of all possible interactions,

$$\mathbf{v}_t = \mathbf{w}_{t-1} \otimes \mathbf{x}_t = (\tilde{w}_{1,t-1}, \tilde{w}_{1,t-1}x_{2,t}, \dots, \tilde{w}_{1,t-1}x_{Q,t}, \tilde{w}_{2,t-1}, \tilde{w}_{2,t-1}x_{2,t}, \dots, \tilde{w}_{2,t-1}x_{Q,t})'$$

(again, recall $x_{1,t} = 1$), where $\tilde{w}_{t,s} = \max(0, w_{t,s})$. We cannot use [eq. \(S2.8\)](#) directly, because it is not $V \times S$. To maintain the linear relationship with parameters in ρ (and, thus, direct sampling), we introduce a sparse matrix. For the $S = 2$ case this sparse coefficient matrix for [eq. \(S2.8\)](#) is

$$\mathbf{P} = \begin{pmatrix} \rho_{11} & 0 \\ \vdots & \vdots \\ \rho_{Q1} & 0 \\ 0 & \rho_{12} \\ \vdots & \vdots \\ 0 & \rho_{Q2} \end{pmatrix} \quad [\text{S2.9}]$$

\mathbf{P} is a $V \times S$ matrix, but it holds at most only QS non-zero elements. For reference purposes the subscripts for ρ in [eq. \(S2.9\)](#) are indicated not by the location in \mathbf{P} , but rather by the species-environment combination each represents, the index (q, s) . The traditional DI growth rate growth rates are given by the elements $\rho_{11}, \rho_{12}, \dots, \rho_{1S}$. This is the first row of matrix ρ , which is multiplied by the intercept in \mathbf{x} , i.e., 1. This is the DI growth rate that is realized in the absence of covariates (or covariates equal to zero). The structure of [eq. \(S2.9\)](#) is imposed to allow for direct (Gibbs) sampling of the coefficients and covariance matrix ([section S7](#)).

The density-dependent (DD) term contains the interactions, i.e., the unique elements of the $S \times S$ matrix α that operates on species pairs, $\tilde{\mathbf{w}}_{t-1} \tilde{\mathbf{w}}'_{t-1}$. These products are held in a length- U vector, where $U \leq S^2$. This vector is limited to interactions that can actually occur. If all species interact with all others, the vector for two species is

$$\mathbf{u}_t = (\tilde{w}_{1,t-1}\tilde{w}_{1,t-1}, \tilde{w}_{1,t-1}\tilde{w}_{2,t-1}, \tilde{w}_{2,t-1}\tilde{w}_{2,t-1})'$$

As for ρ , we need a reorganized version of α that permits direct sampling. Assuming these two species interact, the $U \times S$ matrix of species interaction coefficients is

$$\mathbf{A} = \begin{pmatrix} \alpha_{11} & 0 \\ \alpha_{12} & \alpha_{21} \\ 0 & \alpha_{22} \end{pmatrix} \quad [\text{S2.10}]$$

Again, subscripts indicate not the location in \mathbf{A} , but rather the species combination represented in $\alpha_{s,s'}$. The second row includes both interspecific interactions, because both multiply the product $w_s w_{s'}$ in vector \mathbf{u}_t .

In addition to the equilibrium abundances solved from [eq. \(S2.5\)](#), there is a carrying capacity for the species that is predicted in the absence of movement and regulation by other species. The carrying capacity that applies to the mean environment includes environmental effects is

$$w_s^c | \bar{\mathbf{x}} = -\frac{1}{\alpha_{ss}} \mathbf{P}'_s \bar{\mathbf{x}} \quad [\text{S2.11}]$$

where $\mathbf{P}_s = [\rho_{1,s}, \dots, \rho_{Q,s}]'$ is the non-zero rows of column s in the matrix \mathbf{P} , i.e., only the rows corresponding to species s in eq. (S2.9). Note that this solution for a species in the absence of others can never be non-linear with respect to an environment gradient in \mathbf{x} unless, of course, \mathbf{x} includes quadratic terms. In fact, if predictors in \mathbf{x} are centered and standardized, then all terms in $\bar{\mathbf{x}}$ are zero except the intercept, and eq. (S2.11) collapses to eq. (S2.2).

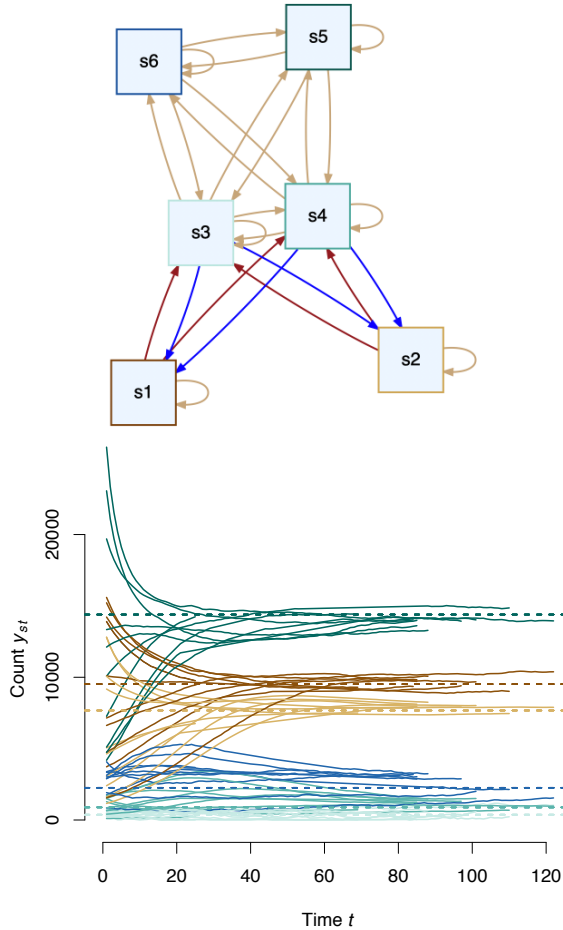


Fig. S2.2. A simulated community of $S = 6$ species on ten sites with $Q = 3$ environmental predictors, started from random initial values, with a mean sampling duration of 100 iterations. Species s_1 and s_2 are predators on s_3 and s_4 (red negative and blue positive arrows). Interspecific competition (brown arrows) occurs between $s_3, s_4, s_5,$ and s_6 . All species experience intraspecific competition. Dashed lines at right show \mathbf{w}^* from eq. (S2.3). The $w_{s,t}$ are latent versions of observed $y_{s,t}$. Colors at right correspond to box outlines at left.

S3. DATA SIMULATION

The simulator makes three important contributions to this analysis. First, it demonstrates that the model can predict the data used to fit the model. This would not be the case if there are errors in the model or in the code that affect either

model fitting or data generation; the model is *generative*, and the simulator confirms it. Second, it demonstrates that, when data are adequate, we can recover parameters used to simulate the data. We would not have this outcome if the model were overfitted*. Finally, it allows us to determine the contributions of food web topology and environment to parameter estimates and data prediction.

The simulator takes as inputs the numbers of species, sites, predictors, and time increments. Each of the three terms in *gjamTime* (movement, DI growth, and DD growth) can be included in a simulation, or not. The topology of the food web is specified in terms of competition and predation, through the signs in the interaction matrix. Observation effort determines how much of each population will be observed. The simulator selects at random parameter values from a range that can produce realistic dynamics. Forward simulation then generates latent $w_{s,t}$ from eq. (S2.5). Based on the specified effort, observed $y_{s,t}$ are generated from eq. (S2.7).

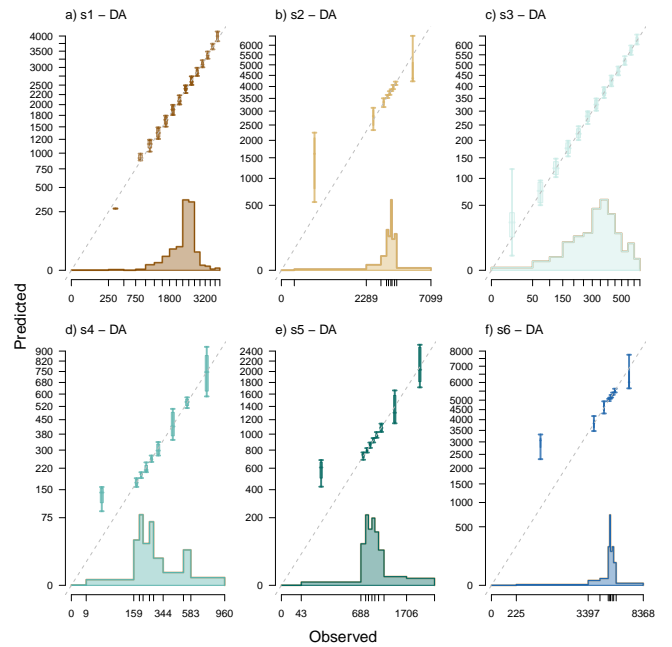


Fig. S3.3. Data prediction for species abundances in the simulated example from fig. 3 of the main text. Colors match time series in fig. S2.2.

The simulator is available at [gjamTimePNASVignette](#), where we use it to show how knowledge of species and environment affect model fitting and prediction. The examples included there step through some simple food webs to emphasize how each of the effects might be examined.

S4. CONTRIBUTIONS TO DYNAMICS

The relative contributions of species-environment interactions through movement, DI growth, and DD are available through the three terms in eq. (S2.5). One way to understand these

*Checks for overfitting with actual data are highly indirect through out-of-sample prediction; this is a desperate measure used because 'true' parameters are unknown. Simulation allows for direct evaluation of overfitting through parameter recovery.

effects is through the increase in residual variance that occurs when a term is left out of the model. For each term in eq. (S2.5) there is a corresponding contribution to overall variance. For environmental effects on movement it is

$$\begin{aligned}\mathbf{c}_E &= \text{diag}(\mathbf{C}_E) \\ C_E &= \text{trace}(\mathbf{C}_E)\end{aligned}\quad [\text{S4.12}]$$

where $\mathbf{C}_E = \beta' \text{Cov}(\mathbf{X}) \beta$, and \mathbf{X} is the $n \times Q$ design matrix having rows \mathbf{x}'_t . The first term in eq. (S2.5) makes a big contribution to dynamics if values in β are large and environmental variation in \mathbf{X} is large. The length- S vector \mathbf{c}_E holds the variance by species, and C_E is the variance across all species.

Species-environmental effects on DI growth contribute

$$\begin{aligned}\mathbf{c}_{EW} &= \text{diag}(\mathbf{C}_{EW}) \\ C_{EW} &= \text{trace}(\mathbf{C}_{EW})\end{aligned}\quad [\text{S4.13}]$$

where $\mathbf{C}_{EW} = \mathbf{P}' \text{Cov}(\mathbf{V}) \mathbf{P}$, and \mathbf{V} is the $n \times V$ design matrix having rows \mathbf{v}'_t . Recall, \mathbf{v}_t holds all combinations of $w_{s,t} x_{q,t}$. $\text{Cov}(\mathbf{V})$ is the covariance for interactions between species abundances and environmental variables. The second term in eq. (S2.5) makes a big contribution if values in \mathbf{P} are large and variation in \mathbf{V} is large.

Species interactions contribute through DD growth,

$$\begin{aligned}\mathbf{c}_{WW} &= \text{diag}(\mathbf{C}_{WW}) \\ C_{WW} &= \text{trace}(\mathbf{C}_{WW})\end{aligned}\quad [\text{S4.14}]$$

where $\mathbf{C}_{WW} = \mathbf{A}' \text{Cov}(\mathbf{U}) \mathbf{A}$, and \mathbf{U} is the $n \times U$ design matrix having rows \mathbf{u}'_t . The $n \times U$ matrix \mathbf{U} has rows that hold all combinations of $w_{s,t} w_{s',t}$. $\text{Cov}(\mathbf{U})$ is the covariance for interactions between species abundances, which depend on all forces affecting the community. This term makes a big contribution if values in \mathbf{A} are large and the variance in species interactions is large.

Omitting these terms from the model inflates the noise in the data from Σ to

$$\Sigma + \mathbf{C}_E + \mathbf{C}_{EW} + \mathbf{C}_{WW} \quad [\text{S4.15}]$$

Thus, each of the ways in which species and environment interact contributes to the fitted model. Leaving them out of the model affects the interpretation depending on each operates through movement, population growth, and/or species interactions.

S5. INTERACTIONS AND INDIRECT EFFECTS

In this section we show that the model captures both non-linear and interaction responses to environmental gradients that are induced through interactions with other species. First we note that, in the absence of species interactions, the carrying capacities for individual species can have non-linear and interaction responses to the environment only if they are specified as direct effects (e.g., as quadratic terms in matrix \mathbf{P} in eq. (S2.11)).

Second, we recognize that any non-linearities or interactions in environment response must engage terms in the model that involve products of predictors, $x_{q,t} \cdot x_{q',t}$ or the species abundances that depend on them, $w_{s,t}(x_{q,t}) \cdot w_{s',t}(x_{q',t})$. An *induced non-linearity* requires one or more of these productions where $q' = q$. An *induced interaction* requires one or more products where $q' \neq q$. The analysis that follows extracts these effects.

The ways in which a predictor q affects the abundance of a species s includes direct and indirect effects. Consider $s = 1, \dots, S$ species and $q = 1, \dots, Q$ environmental variables. The abundance of species s changes by increment $\Delta w_{s,t}$ as $w_{s,t+1} = w_{s,t} + \Delta w_{s,t}$. The effect of a variable q on growth of species s can depend on other species s' ,

$$\begin{aligned}\frac{d\Delta w_{s,t}}{dx_{q,t}} &= \frac{\partial \Delta w_{s,t}}{\partial x_{q,t}} + \sum_{s' \neq s} \frac{\partial \Delta w_{s,t}}{\partial w_{s',t}} \frac{\partial w_{s',t}}{\partial x_{q,t}} \\ &= \text{direct effect} + \text{indirect/interactions}\end{aligned}\quad [\text{S5.16}]$$

The first term is direct, coming through the movement and DI growth terms of eq. (S2.5). If there are interactions in the model specified as direct effects, then the first term is

$$\begin{aligned}\frac{\partial \Delta w_{s,t}}{\partial x_{q,t}} &= \beta_{q,s} + \sum_{\{q'\}} \beta_{qq',s} \mathbf{x}'_{\{q'\},t} + w_{s,t} \left[\rho_{q,s} + \sum_{\{q''\}} \rho_{qq'',s} \mathbf{x}'_{\{q''\},t} \right] \\ &= \beta_{q,s} + \boldsymbol{\beta}_{\{q'\}s'} \mathbf{x}'_{\{q'\},t} + w_{s,t} \left[\rho_{q,s} + \boldsymbol{\rho}_{\{q''\}s'} \mathbf{x}'_{\{q''\},t} \right]\end{aligned}$$

where $\{q'\}$ is the set of movement predictors that interact with q , $\beta_{s,qq'}$ is the interaction effect of q and q' on movement, $\{q''\}$ is set of DI growth predictors that interact with q , and $\rho_{qq'',s}$ is the interaction effect of q and q'' on DI growth. If the model does not specify interactions, then this reduces to

$$\frac{\partial \Delta w_{s,t}}{\partial x_{q,t}} = \beta_{q,s} + w_{s,t} \rho_{q,s} \quad [\text{S5.17}]$$

Because we are concerned with interactions that are induced through other species, we refer to this simplified expression for the remainder of the analysis.

The second term of eq. (S5.16) is indirect, coming through environment effects on the other species. The first factor in the summation contributes a length- $(S-1)$ vector

$$\begin{aligned}\mathbf{c}_{-s,t} &= \mathbf{w}'_{-s,t} \boldsymbol{\alpha}'_{-s} + \Sigma_{s,s'} \boldsymbol{\Sigma}_{s',s'}^{-1} \\ &= \text{species interactions} + \text{indirect effect}\end{aligned}\quad [\text{S5.18}]$$

where $\boldsymbol{\alpha}_{-s}$ omits column s of matrix $\boldsymbol{\alpha}$ (it holds the effects of all other species on s), and $\mathbf{w}_{-s,t}$ is the vector of other species abundances. Indirect effects also come conditionally through the ways in which species covary that are not accounted for in the mean structure of the model, held in the residual covariance matrix Σ . This second term cannot depend on x_q so it cannot contribute to non-linear responses; we hereafter omit it.

With these results we can identify the sources of non-linearities and interactions. Because these can only come

through products involving $x_{q,t}$, we focus on the vector $\mathbf{w}_{s,t}$, which depends on $x_{q,t}$ through crossproduct terms $w_{s,t}(x_{q,t}) \cdot w_{s',t}(x_{q',t})$. The terms that remain from eq. (S5.16) after excluding those that cannot account for non-linearities:

$$\sum_{s' \neq s} \frac{\partial \Delta w_{s,t}}{\partial w_{s',t}} \frac{\partial w_{s',t}}{\partial x_{q,t}} \propto w_{s,t} \boldsymbol{\alpha}_{s,-s} (\boldsymbol{\rho}_{q,-s} \circ \mathbf{w}_{-s,t}) \quad [\text{S5.19}]$$

where $\boldsymbol{\alpha}_{s,-s}$ is row s of $\boldsymbol{\alpha}_{-s}$, $\boldsymbol{\rho}_{q,-s}$ is a length- $(S-1)$ vector of elements $\{\rho_{q,s'}\}_{s' \neq s}$, and \circ is the Hadamard product. The form of this expression makes it clear that products of predictors $x_{q,t} x_{q',t}$ can only come through their effects on products of species abundances $w_{s,t} w_{s',t}$. The magnitudes of these indirect effects depend on coefficients in $\boldsymbol{\alpha}$ and $\boldsymbol{\rho}$. These effects are quadratic (non-linear) when two species depend on the same predictor, $w_{s,t}(x_{q,t}) \cdot w_{s',t}(x_{q,t})$. They are interactive when species that interact, through $\alpha_{s,s'} \neq 0$, depend on different predictors $w_{s,t}(x_{q,t}) \cdot w_{s',t}(x_{q',t})$.

Taken together, the terms in the model have no way of generating non-linear effects, but through the densities of others. If all other species are absent, then eq. (S5.19) is equal to zero. Even where all terms in the model are linear for all predictors, non-linear gradient responses are 'induced' by the responses of other species with which a species interacts.

S6. PRIOR KNOWLEDGE

Food web theory is used *a priori* to reduce dimensionality. For transparency, prior distributions are uniform and bounded over intervals based on ecological understanding. Where no information exists, a non-informative prior can be uniform over the real line, $(-\infty, \infty)$. Informative priors can be truncated at zero, to reflect knowledge of the signs of effects and interactions (section S6). To set the scale for coefficients, an approximate range of growth rates starts with abundance per effort, $w_{s,t} \approx \frac{y_{s,t}}{E_{s,t}}$. Realistic bounds of possible rates of change can be combined with liberal estimates of how movement, DI growth, and DD group could contribute to this change. Insight comes from basic demographic rates, which can involve knowledge of dispersal potential, numbers of litters per year, clutch sizes, survival rates, and so on. This information is used only to set liberal bounds on possible rates. For DI growth in \mathbf{P} , there are ESI that contribute to the growth rate as products of $\tilde{w}_{s,t} \cdot x_{q,t}$,

$$\mathbf{P}'_s \mathbf{v}_t = \rho_{1s} \tilde{w}_{s,t} + \sum_{q=2}^Q \rho_{qs} \tilde{w}_{s,t} x_{q,t} \quad [\text{S6.20}]$$

where \mathbf{P}_s is a column s of $V \times S$ matrix \mathbf{P} . If predictors in \mathbf{x}_t are centered, the parameters ρ_{1s} are the DI growth rate for each species at mean covariate values in eq. (S2.9). They must be greater than 0 for net growth to occur, and vice versa. A deviation from ± 0.1 would contribute to a 10% decline or increase per time increment. For a population capable of DI growth of 1% per year, we expect values in the range of 0.01 yr^{-1} . For a population capable of doubling in 10

hr, we expect values in the range of 0.07 hr^{-1} . Populations are also capable of decline, values of $\rho_{1s} < 0$. However, a population that experiences competition that is not reliant on immigration (it is not a 'sink population') can be expected to have a positive growth rate ($\rho_{1s} > 0$ or, equivalently, $\mathbf{P}_{0s} > 0$). This knowledge is the basis for a prior interval on the ρ_{1s} coefficients, which are the intercept terms in eq. (S2.9). For the BBS example, our prior interval for ρ_{1s} considered the range of clutch sizes expected for species in this example and combined it with juvenile survival to define a broad prior interval $(-0.05, 0.1)$, which ranges from 50% decline in 13 years to population doubling in 7 years, i.e., much wider than would be expected from simple DI growth at mean covariate values. Non-intercept coefficients in \mathbf{P} apply to standardized covariates in \mathbf{x} , which have mean zero, multiplied by $w_{s,t}$. The non-intercept coefficients $\rho_{2s}, \dots, \rho_{Qs}$ are also liberally bounded based on possible ranges for DI growth under the unit-variance range of \mathbf{x} .

The movement coefficient matrix holds effects that depends on the growth rate and the variation in a predictor q ,

$$\beta_{qs} = \frac{\partial}{\partial x_q} \left(\frac{dw_s}{dt} \right) \quad [\text{S6.21}]$$

Possible values for this coefficient consider the range of growth rates and the range of x_q . The range of variation in dw_s/dt that could occur due to movement can, in many cases, come from existing knowledge. The intercepts in $\boldsymbol{\beta}_{1s}, s = 1, \dots, S$ describe movement that is not linked to predictors in \mathbf{x} . For the WEL example, we assumed no movement at all, so the first term of eq. (S2.5) is omitted. This assumption does not exclude the possibility that a small amount of movement might occur, only that it would be too small to estimate from the data. For the coefficients that are not intercepts we consider variation in predictors. We standardize predictors in \mathbf{x} , so we can think about effects having fluctuations over this range of variation (e.g., several standard deviations for x_q).

Flat prior intervals for DD growth in \mathbf{A} might be assessed in several ways. We know that many species do not directly interact. These species pairs are zero in $\boldsymbol{\alpha}$ and in the corresponding elements in \mathbf{A} . This does not mean that there are no indirect interactions between these species pairs, only that those indirect effects must enter through $\boldsymbol{\Sigma}$ or through lagged effects $\alpha_{s's''} w_{s'',t-1} \rightarrow \alpha_{ss'} w_{s',t} \rightarrow \Delta w_{s,t}$. The effects in matrix $\boldsymbol{\alpha}$ are non-lagged and direct. Where we have less information, the signs of elements $\alpha_{s,s'}$ may be known from trophic relationships. Liberal bounds on $\alpha_{s,s'}$ are based on a $S \times S$ matrix of rates and crossproducts, $\frac{dw_s}{w_s w_{s'}}$, i.e., the range of population growth rates and the magnitudes of the products that determine their contributions to growth.

S7. POSTERIOR SIMULATION

A. Sampling sparse coefficient matrices. The multivariate variable selection literature focuses primarily on selecting entire rows in coefficient matrices like $\boldsymbol{\beta}$, corresponding to columns in the design matrix (Brown et al. 1998, Lique et al. 2016). Here we are not attempting to eliminate whole rows of $\boldsymbol{\beta}$,

\mathbf{P} , and \mathbf{A} , but rather we sample conditional on zeros that can be specified anywhere in these matrices. Again, we are not selecting variables as part of model fitting, but rather specifying the combinations that can be non-zero *a priori*. This is equivalent to identifying the species that should not be linked in figure 2 of the main text. The strategy that follows recognizes that coefficient matrices can be large, but also sparse. Recall that much of the sparsity comes from the need to sample coefficients directly, which is demonstrated in this section.

We begin by stacking the row vectors $\Delta \mathbf{w}'_{i,t}$, $\mathbf{x}'_{i,t}$, $\mathbf{v}'_{i,t}$, and $\mathbf{u}'_{i,t}$ to generate $m \times S$ matrix \mathbf{W} , $m \times Q$ matrix \mathbf{X} , $m \times V$ matrix \mathbf{V} , and $m \times U$ matrix \mathbf{U} , where $m = \sum_i^n T_i$. The joint distribution is now matrix-normal

$$MN_{m,S}(\mathbf{W}|\mathbf{X}\beta + \mathbf{V}\mathbf{P} + \mathbf{U}\mathbf{A}, \mathbf{I}_m, \Sigma)$$

where \mathbf{I}_m is the identity matrix.

Sampling the three coefficient matrices can be done in the same way. Using the example of matrix \mathbf{A} , we have the conditional posterior distribution

$$[\mathbf{A}|\beta, \mathbf{P}, \Sigma] = MVN_{SU}(\text{vec}(\mathbf{A})|\mathbf{m}, \mathbf{M})$$

where $\text{vec}(\mathbf{A})$ stacks the columns of \mathbf{A} into a single vector. Sampling the full matrix starts from the usual conditional mean and covariance matrices,

$$\begin{aligned} \mathbf{m} &= \text{vec} \left((\mathbf{U}'\mathbf{U})^{-1} \mathbf{U} (\mathbf{W} - \mathbf{X}\beta - \mathbf{V}\mathbf{P})' \right) \\ \mathbf{M} &= (\mathbf{U}'\mathbf{U})^{-1} \otimes \Sigma \end{aligned} \quad [\text{S7.22}]$$

The matrix \mathbf{A} includes elements that must be sampled, designated \mathbf{A}_c , and others fixed at zero, designated $\mathbf{A}_{c'}$, where c holds locations of non-zero elements in $\text{vec}(\mathbf{A})$, and c' holds the locations of zero elements in $\text{vec}(\mathbf{A})$.

We need to condition on zeros, which can be done in a way that reduces dimensionality. Matrix \mathbf{A} is vectorized and ordered with elements c followed by c' , such that

$$\mathbf{m} = \begin{pmatrix} \mathbf{m}_c \\ \mathbf{m}_{c'} \end{pmatrix}, \mathbf{M} = \begin{pmatrix} \mathbf{M}_{cc} & \mathbf{M}_{cc'} \\ \mathbf{M}_{c'c} & \mathbf{M}_{c'c'} \end{pmatrix}$$

We want to sample non-zero elements having conditional mean and covariance

$$\begin{aligned} \mathbf{m}_{c|c'} &= \mathbf{m}_c - \mathbf{M}_{cc'} \mathbf{M}_{c'c'}^{-1} \mathbf{m}_{c'} \\ \mathbf{M}_{c|c'} &= \mathbf{M}_{cc} - \mathbf{M}_{cc'} \mathbf{M}_{c'c'}^{-1} \mathbf{M}_{c'c} \end{aligned} \quad [\text{S7.23}]$$

The problem here is that the inverse $\mathbf{M}_{c'c'}^{-1}$ on the right side of eq. (S7.23) is too large, because there can be many species, and most elements in \mathbf{A} could be zero ($c' \gg c$).

Exploiting properties of the Kronecker product, the conditional covariance matrix is available from

$$\mathbf{M}_{c|c'} = \mathbf{C}_{cc}^{-1}$$

where \mathbf{C}_{cc} is a matrix having elements $\mathbf{C}_{cc[ij]} = [\Sigma^{-1} \otimes \mathbf{U}'\mathbf{U}]_{ij}$ for rows i and columns j (note reverse order for Kronecker from eq. (S7.22)). This avoids inverting the matrix $\mathbf{M}_{c'c'}$ in the second line of eq. (S7.23). However, this inverse is still needed for the conditional mean in the first line of eq. (S7.23). In contrast to $\mathbf{M}_{c'c'}$ the matrix \mathbf{C}_{cc} will be small and provides access to the inverse through

$$\mathbf{M}_{c'c'}^{-1} = \mathbf{C}_{c'c'} - \mathbf{C}_{c'c} \mathbf{C}_{cc}^{-1} \mathbf{C}_{cc'}$$

In other words, we invert the small matrix \mathbf{C}_{cc} to obtain the inverse of the big, sparse matrix $\mathbf{M}_{c'c'}$. Note that for dimension reduction described in (22) we also need a conditional covariance, which in that case simplifies to $\mathbf{M}_{c|c'} = \sigma^2 [(\mathbf{U}'\mathbf{U})_{cc}]^{-1}$. Albeit circuitous, this method is fast and accommodates large numbers of species.

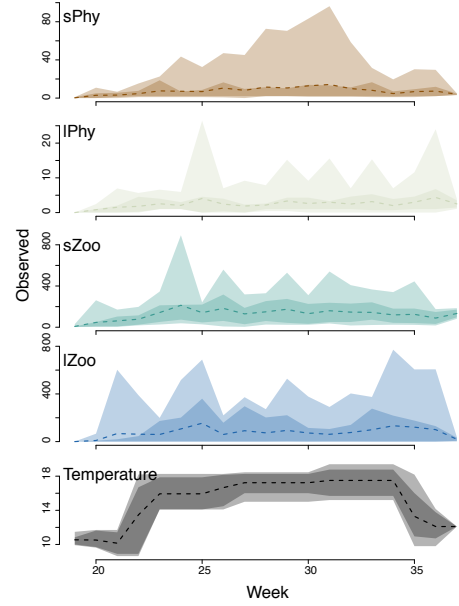


Fig. S7.4. The ranges of observed values for four species groups and across three lakes and six years by week of the year with temperature (below). Shading bounds 68% and 95% of observations, respectively.

B. Sampling latent states. This non-linear model does not admit direct sampling of states $w_{s,t}$. These are sampled with a Metropolis step. We propose from

$$w_{s,t}^* \sim N(w_{s,t}, v_{s,t}) \times 1(p_{s,k} < w_{s,t}^* \leq p_{s,k+1})$$

where $v_{s,t}$ is a proposal variance that will adapt based on acceptance rates and posterior variance, and $(p_{s,k}, p_{s,k+1}]$ is the partition. Acceptance is based on

$$\begin{aligned} [\mathbf{w}_t^* | \mathbf{w}_{t-1}, \mathbf{w}_{t+1}] &\propto MVN(\Delta \mathbf{w}_t^* | \beta' \mathbf{x}_t + \mathbf{P}' \mathbf{v}_t + \mathbf{A}' \mathbf{u}_t, \Sigma) \\ &\times MVN(\Delta \mathbf{w}_{t+1} | \beta \mathbf{x}_{t+1} + \mathbf{P} \mathbf{v}_{t+1}^* + \mathbf{A} \mathbf{u}_{t+1}^*, \Sigma) \end{aligned}$$

For time zero, only the second factor is used.

Table S7.2. Summary of lakes used in the Wisconsin Experimental Lakes example.

Lake ^a	Treatment ^b	Effects
Paul Lake (bass-dominated)	reference	–
Peter Lake (high planktivory)	nutrients, bass removed	increased planktivory, decreased <i>D. pulex</i> , increased phytoplankton
West Long Lake (bass-dominated, low planktivory)	nutrients	reduced planktivory, increased <i>D. pulex</i> , suppressed phytoplankton response

^a Lake characteristics summarized in parentheses.

^b Summary of manipulation

S8. APPLICATIONS TO FOOD WEB DATA

A. Wisconsin Experimental Lakes. Experimental manipulation of nutrients and largemouth bass in three lakes was implemented to determine how top-down controls by piscivores could mediate the effects of nutrient loading (table S7.2). Removal of piscivorous bass from Peter Lake in 1991 was expected to increase population growth rates of phytoplankton. This would occur through reduced largemouth bass predation on the smaller fish (i.e., *piscivory*) that consume the dominant grazer *D. pulex*; these smaller fish are *planktivorous*. Under decreased piscivory, planktivorous fish increase, thereby reducing grazing pressure by *D. pulex*. This reduction in grazing pressure allows phytoplankton blooms under high nutrient loading (23). Lakes were monitored for two pre-treatment years (1991, 1992) and four subsequent years of nutrient addition (table S7.2). Mean observed values over years are in fig. S7.4. Data collection and trends for each of three lakes are detailed in the original publications.

Data for this example come from (2) for four species groups and the two additional variables (bass) and nutrient addition (Pvol). As in previous analyses (2, 24) we treat nutrient addition (pVol) and planktivory (bass: interpolated planktivore abundances from observations at the beginning and end of each year) as external controls on the community defined by the aggregate four species groups. Our fitted model included terms for DI and DD growth (ρ , α), but not movement. It includes an interaction between bass and temperature (temp).

The sample contains $n = 291$ observations on $S = 4$ response variables (species groups). To allow for observation error, continuous variables were discretized to integer values; if this example were not for demonstration only, this discretization would benefit from prior understanding of variability in observations. Each lake-year is treated as a separate time series (2). The time step is 2 weeks. Each lake-year has a prior value for y_{i0} that is based on subsequent values. This value is not fixed—it has a corresponding posterior distribution. Computation involved 80,000 Gibbs steps, with a burnin of 40,000.

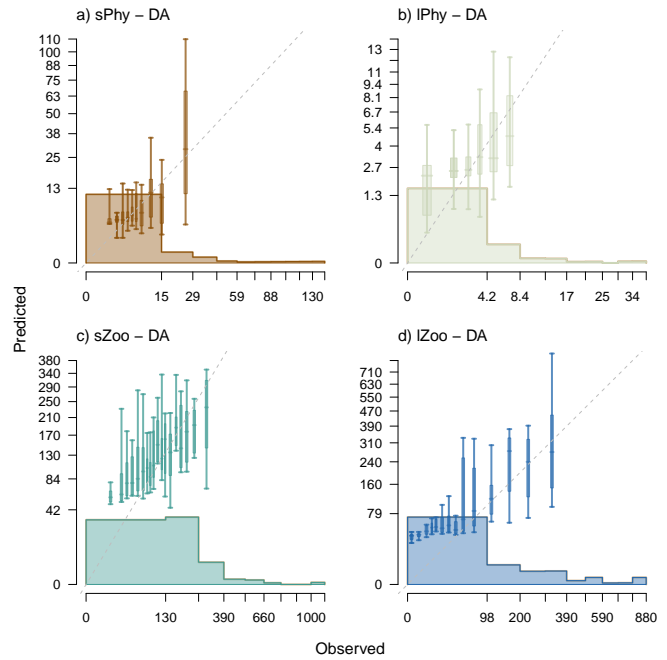


Fig. S8.5. Predicted and observed observations from the fitted model by species group. Boxes and whiskers bound 68% and 95% of observations, respectively. The basal histogram shows the distribution of data.

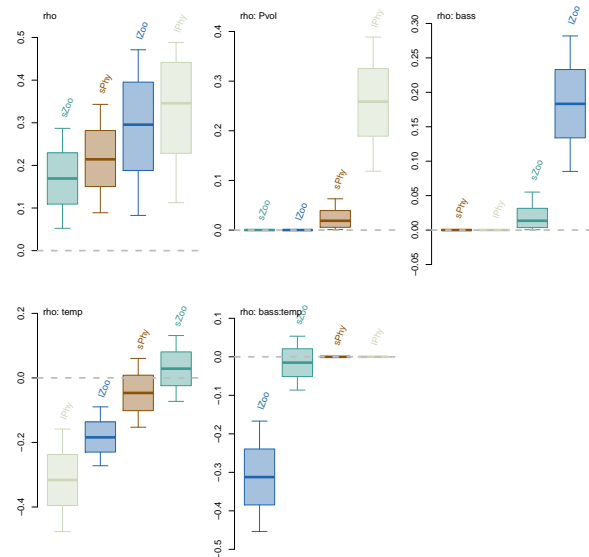


Fig. S8.6. Posterior parameter estimates for DI growth responses to environmental variables in matrix ρ for the WEL example. Boxes and whiskers bound 68% and 95% of the posterior distribution, respectively.

The model predicts the data (fig. S8.5) and it resolves the posterior. The RMSPE is 92.8, and the DIC is 39,574. Coefficients in DI growth matrix ρ are shown as boxplots in fig. S8.6. Parameter estimates quantify main effects and an interaction between bass and temp. Comparisons of prior and posterior for ρ and interaction matrix α show that data have updated the prior distribution (fig. S8.7). DI growth rates at mean covariate levels ('intercept' in fig. S8.6) range from near zero to 0.5 wk^{-1} , the latter corresponding to a doubling time of $< 2 \text{ wk}$. Judging from prior-posterior comparisons, these

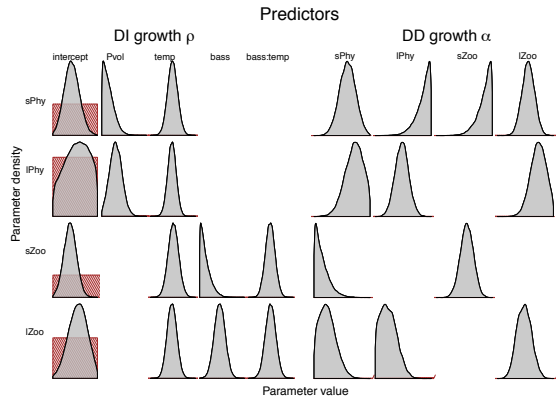


Fig. S8.7. Bayesian learning in the WEL application is evident from a comparison of prior (brown rectangles) and posterior (grey densities). The prior is not visible for parameters where data dominate. Missing combinations (white spaces) are zero *a priori*. Only intercepts in ρ (the DI growth rate at mean values for covariates, shown at left) are weakly identified.

are the hardest parameters to identify (fig. S8.7). Nutrient addition (Pvol) has a stronger direct effect on large than small phytoplankton (fig. S8.6). By depressing planktivory, piscivorous bass had a strong positive effect on zooplankton, especially *D. pulex* (lZoo).

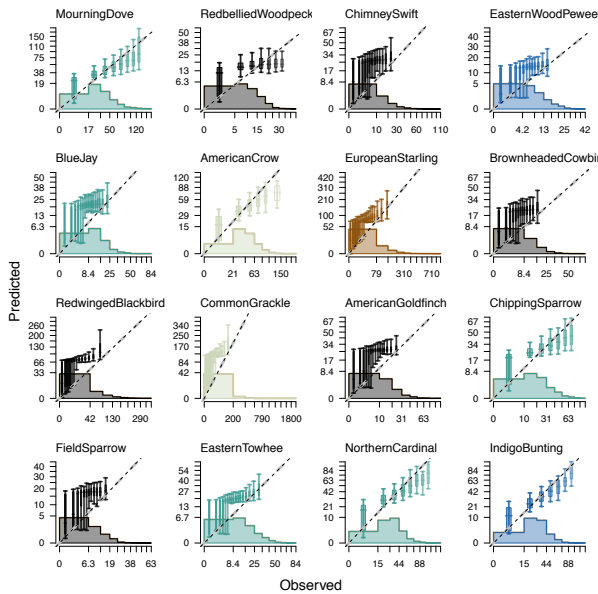


Fig. S8.8. Predicted and observed observations from the fitted model for the first 16 species in BBS. Boxes and whiskers bound 68% and 95% of observations, respectively. The basal histogram shows the distribution of data.

Eigenvalues of $\hat{\alpha}$ indicate damped oscillations, having all negative real parts (locally stable) and non-zero imaginary parts (the expected oscillations for this food web that includes predation), $(-0.0386, -0.00136 \pm 0.00350i, -0.00103)$. The analysis is reproduced at [gjamTimePNASVignette](#).

B. Breeding bird survey. For this example, we limited BBS data to a region centered on mid-Atlantic states of eastern USA between longitude $(-86^\circ, -77^\circ \text{ W})$ and latitude $(34^\circ, 46^\circ \text{ N})$ and spans a range of climate and land cover types (fig. 5).

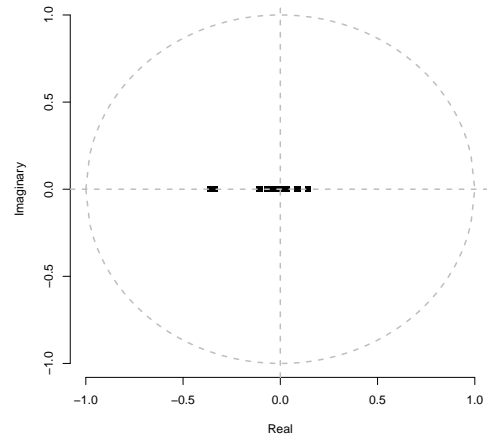


Fig. S8.9. Many eigenvalues for α in the BBS example are positive. None have imaginary parts.

Because survey coverage remained unstable in early years, we limited analysis to post-1996 observations. To simplify output we further limited the analysis to species present in at least 6000 route-years. This reduced data set with few predictors was used for purposes of transparency; the approach will accommodate many more species and observations. The matrix β was limited to wind conditions at the time of observation, which can affect activity and, thus, movement. Wind is a four-level factor. The matrix ρ was limited to land cover type from the National Land Cover Database taken at the location of the starting point for each BBS transect. This index is crude, because transects span 25 km and, thus, multiple land cover types, which are not located in the BBS data. NLCD types were aggregated into a six-level factor as shown in figure 5. Additional predictors in ρ are June temperature and summer moisture deficit (P - PET). The α matrix was sparsified based on liberally defined guilds of potential competitors that were defined by diet, foraging habits, and body size (fig. 2B).

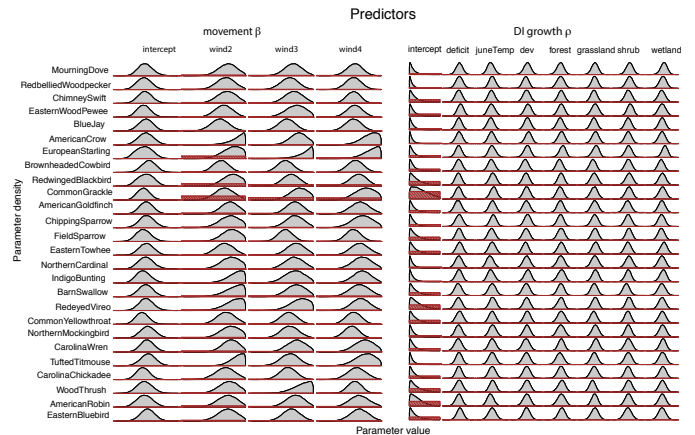


Fig. S8.10. Bayesian learning in the BBS application is evident from a comparison of prior (brown rectangles) and posterior (grey densities) for movement matrix β and DI growth matrix ρ . The prior is not visible for panels where data dominate. Only intercepts in ρ (the DI growth rate at mean values for covariates) are weakly identified.

The sample contains $n = 7924$ observations on $S = 26$

species for 206,024 observed counts. Counts per effort range from 0 to 20.2. Missing values come from the imputed initial abundance for each BBS route and additional route-years in which counts were not available. There are 6060 missing values in \mathbf{X} and 35,906 missing values in \mathbf{Y} . Imputation of missing values is based on a prior mean of adjacent values in the time series and a weak prior weight (one tenth of the effort for non-missing values). The RMSPE for the fitted model is 50.4, and the DIC is 1,214,349. Computation involved 50,000 Gibbs steps, with a burnin of 20,000.

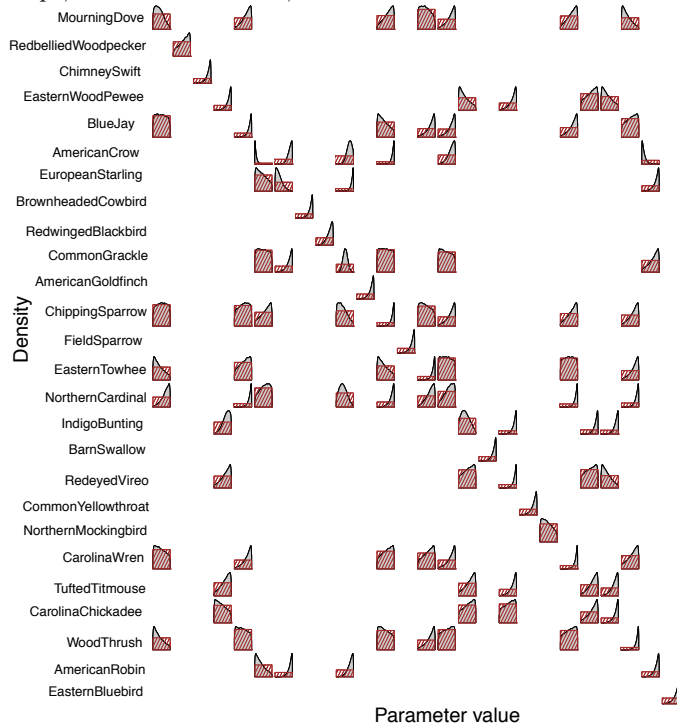


Fig. S8.11. Bayesian learning in the BBS application is evident from a comparison of prior (brown rectangles) and posterior (shaded densities) interaction matrix α . Columns are ordered the same as rows. Intraspecific competition is represented by the diagonal, from upper left to lower right. Prior and posterior are essentially identical where data do not update the prior. Missing combinations (white spaces) are zero a priori.

Included in [gjamTimePNASVignette](#) are posterior estimates of variables affecting DI and DD growth rates that contribute to the estimates and predictions in figures 4 and 5. Eigenvalues of α include positive values and no imaginary parts (fig. S8.9), indicating that equilibrium estimates are unstable and thus expected to include zero abundances for some species in some locations as is evident from gradients shown in figure 5.

There are two covariates in ρ , June temperature (juneTemp) and moisture deficit. Land-cover types include crop, which is absorbed into the intercept (it is the reference factor level), developed (dev), forest, grassland, shrub, and wetland (fig. 5). The large numbers of coefficients in these tables are given a visual context in [fig. S8.10](#).

2. Ives AR, Dennis B, Cottingham KL, Carpenter SR (2003) Estimating community stability and ecological interactions from time-series data. *Ecological Monographs* 73(2):301–330.
3. Carrara F, Giometto A, Seymour M, Rinaldo A, Altermatt F (2015) Inferring species interactions in ecological communities: a comparison of methods at different levels of complexity. *Methods in Ecology and Evolution* 6(8):895–906.
4. Wikle CK, Milliff RF, Nychka D, Berliner LM (2001) Spatiotemporal hierarchical bayesian modeling: Tropical ocean surface winds. *Journal of the American Statistical Association* 96(454):382–397.
5. Clark J, Bjornstad O (2004) Population time series: Process variability, observation errors, missing values, lags, and hidden states. *Ecology* 85(11):3140–3150.
6. Wikle CK (2003) Hierarchical bayesian models for predicting the spread of ecological processes. *Ecology* 84(6):1382–1394.
7. Inoue A, Kilian L (2013) Inference on impulse response functions in structural var models. *Journal of Econometrics* 177(1):1–13.
8. Ye H, et al. (2015) Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proceedings of the National Academy of Sciences* 112(13):E1569–E1576.
9. Brunton SL, Proctor JL, Kutz JN (2016) Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* 113(15):3932–3937.
10. Schliep EM, et al. (2018) Joint species distribution modelling for spatio-temporal occurrence and ordinal abundance data. *Global Ecology and Biogeography* 27(1):142–155.
11. Ovaskainen O, et al. (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(1855):20170768.
12. Hooten MB, Wikle CK (2008) A hierarchical bayesian non-linear spatio-temporal model for the spread of invasive species with application to the eurasian collared-dove. *Environmental and Ecological Statistics* 15(1):59–70.
13. Chib S, Greenberg E (1998) Analysis of multivariate probit models. *Biometrika* 85(2):347–361.
14. Lawrence E, Bingham D, Liu CH, Nair VN (2008) Bayesian inference for multivariate ordinal data using parameter expansion. *Technometrics* 50(2):182–191.
15. Cressie NAC, Wikle CK (2011) *Statistics for spatio-temporal data*, Wiley series in probability and statistics. (Wiley, Hoboken, N.J.), pp. xxii, 588 p.
16. May R (2001) *Stability and Complexity in Model Ecosystems*. (Princeton University Press, Princeton).
17. Boit A, Martinez ND, Williams RJ, Gaedke U (2012) Mechanistic theory and modelling of complex food-web dynamics in lake constance. *Ecol Lett* 15(6):594–602.
18. Thompson RM, et al. (2012) Food webs: reconciling the structure and function of biodiversity. *Trends Ecol Evol* 27(12):689–97.
19. Jonsson T, Kaartinen R, Jonsson M, Bommarco R (2018) Predictive power of food web models based on body size decreases with trophic complexity. *Ecol Lett* 21(5):702–712.
20. Allesina S, Tang S (2012) Stability criteria for complex ecosystems. *Nature* 483(7388):205–8.
21. Neill WE (1974) The community matrix and interdependence of the competition coefficients. *The American Naturalist* 108(962):399–408.
22. Taylor-Rodriguez D, Kaufeld K, Schliep EM, Clark JS, Gelfand AE (2017) Joint species distribution modeling: Dimension reduction using dirichlet processes. *Bayesian Analysis* 12(4):939–967.
23. Shapiro J (1990) Biomanipulation - the next phase - making it stable. *Hydrobiologia* 200:13–27.
24. Carpenter SR, et al. (2011) Early warnings of regime shifts: A whole-ecosystem experiment. *Science* 332(6033):1079–1082.

1. Clark JS, Nemergut D, Seyedsnasrollah B, Turner PJ, Zhang S (2017) Generalized joint attribute modeling for biodiversity analysis: median-zero, multivariate, multifarious data. *Ecological Monographs* 87(1):34–56.