Ecological examples of nonstationarity, nonlinearity, and statistical interactions in dynamic structural equation models

Running header: nonlinear and nonstationary structural equations

James T. Thorson[1*], Kasper Kristensen[2]

[1] Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, National Marine Fisheries Service

[2] Technical University of Denmark, Lyngby, Denmark

* Corresponding author: James.Thorson@noaa.gov

Authorship statement

J. Thorson developed code in *dsem*, and K. Kristensen modified TMB to enable efficient estimation. J. Thorson led analysis, data curation, and writing, and both authors contributed to editing.

Acknowledgements

We thank Steve Munch for providing the copy of Didinium-Paramesium abundance, Mahaffy (2009) for the digitized copy of Hare-Lynx abundance used here, Daniel Schindler for curating the Lake Washington sampling data, and the MARSS developers (Eli Holmes, Eric Ward, and Mark Scheuerell) for distributing the Lake Washington sampling data in easy-to-use format. We also thank Grant Adams and Jennifer Bigman for helpful comments on an earlier draft.

Data accessibility and reproducibility

Data for sea surface temperature at Departure Bay are available online: https://open.canada.ca/data/en/dataset/719955f2-bf8e-44f7-bc26-6bd623e82884/resource/17c30115-25de-4bad-9286-51d3ef467793, and we use the copy from July 23, 2025 and downloaded Nov. 3, 2025. Data for the Pacific Decadal Oscillation were downloaded from JISAO (http://research.jisao.washington.edu/pdo/PDO.latest.txt) on Nov. 6, 2018, as analyzed previously by Thorson et al. (2020). Data for Lake Washington are from the MARSS package release 3.11.9 (Holmes et al., 2012), as collected by Dr. W. T. Edmondson, funded by the Andrew Mellon Foundation, and curated by Dr. Daniel Schindler.

All analysis is conducted using R-package *dsem* release 2.0.0 (https://github.com/James-Thorson-NOAA/dsem@dev). Code and data to reproduce all figures is available in a GitHub repo (https://github.com/James-Thorson/DSEM-varying-paths), publicly available upon acceptance and distributed as a ZIP file during peer review.

Abstract:

1. Ecologists are adapting structural causal modelling for spatial, phylogenetic, and time-series analysis. However, ecological extensions of path analysis and structural equation models (SEM) typically assume that interactions ("path coefficients") are stationary, linear, and additive, while ecological and evolutionary dynamics are often nonstationary, nonlinear, and include statistical interactions.
2. Here, we combine moderated SEM (estimating path coefficients as model variables) with dynamic SEM (estimating both simultaneous and lagged interactions among variables), develop a new "path-lag-slope" notation to specify this combination, and demonstrate it using a simulation experiment and three ecological case studies.
3. The simulation experiment confirms that an autocorrelated "random-slope" model can estimate the nonstationary impact of one variable on another, but that the random slope is shrunk towards a constant value as data become less informative. The first case study then demonstrates nonstationarity by estimating an autoregressive slope linking a regional climate index to local ocean temperature near Vancouver Island. The second demonstrates nonlinearity by approximating Lotka-Volterra dynamics for two predator-prey systems, which closely match estimates of interactions and carrying capacity from traditional ordinary-differential equation methods. The third demonstrates statistical interactions by using monthly plankton samples (1962-1994) to show that resource-consumer-predator interactions in Lake Washington have a dome-shaped response to temperature.
4. We envision several uses in causal analysis: (1) testing whether path coefficients are nonstationary; (2) estimating nonlinear responses given missing data; and (3) linking ecological parameters to hypothesized drivers in applied modelling.

Keywords: Dynamic structural equation model; random slopes; moderated structural equation model; nonstationary; Lotka-Volterra model

Introduction

Ecologists are showing an increasing interest in causal analysis (Arif & MacNeil, 2022; Byrnes & Dee, 2025; Grace, 2024; Larsen et al., 2019). Ecological applications of causal analysis have typically used path analysis (PA) or structural equation models (SEM) to identify direct and indirect consequences of hypothetical policy or environmental changes. These models require hypothesizing linear structural relationships among a set of variables (e.g., A causes B and B causes C), where the strength of relationships is then estimated from the covariance of samples. The fitted model can be used to identify how a policy changing one variable results in direct and indirect effects on other model variables (in the previous example, for example, a change in B would affect C but not A, despite all variables being correlated).

In particular, ecologists are adapting SEM and PA for ecological contexts such phylogenetic, spatial, and time-series analysis. For example, evolutionary ecologists are using phylogenetic structural equation models (PSEM) or path analysis (PPA) to identify evolutionary relationships among species traits (Thorson et al., 2023; von Hardenberg & Gonzalez-Voyer, 2013). Similarly, community ecologists are developing spatial extensions of SEM to estimate species interactions that underlie observed covariance in densities among species across sites (Papadogeorgou et al., 2023; Thorson et al., 2025). Finally, population ecologists are extending prior developments in multivariate autoregressive (MAR) models (Ives et al., 2003; Wootton & Emmerson, 2005) to develop dynamic structural equation models (DSEM), which estimates both simultaneous and lagged interactions using ecological time-series (Thorson et al., 2024), based on similar models in psychology (Asparouhov et al., 2018).

Despite this growing interests, SEM and PA typically require assuming that ecological relationships are stationary, linear, and additive. These assumptions conflict with the general recognition that many parameters of ecological models vary across time, space, and phylogenies (Ives, 2022; Rollinson et al., 2021), that ecological relationships (e.g., metabolic response to temperature) are nonlinear (Munch et al., 2018), and that species interactions might be context-dependent such that regression models should include statistical interactions (Hixon & Carr, 1997). Existing extensions of SEM and PA can account for nonstationarity, nonlinearity, and statistical interactions in some special cases. For example, when causal relationships are unidirectional and data are free of missing values, ecologists can use generalized additive models within "piecewise SEM" (Lefcheck, 2016) to estimate nonlinear linkages. Similarly, SEM (and its spatial, temporal, and phylogenetic extensions) can incorporate latent variables to identify some forms of nonstationarity (i.e., random covariation in average densities among sites). Alternatively, SEM can be extended to incorporate "moderated" interactions, e.g., where the effect of covariate X on response Y depends upon a moderating variable W (i.e., $\mathbb{E}(Y) = \beta_X X + \beta_{XW} XW$). The moderated variable W might be observed (such that XW is included as covariate to represent the statistical interaction of X and W) or unobserved (such that $\beta_{XW} W$ is treated as a random-slope for covariate X). In either case, a "moderated-SEM" (MSEM) can be applied to correct for the statistical interaction or nonstationary slope parameter. However, MSEM was not discussed in recent ecological reviews for causal modelling (Arif & MacNeil, 2022; Byrnes & Dee, 2025; Grace, 2024; Larsen et al., 2019), and has seen limited use in ecological modelling (although see Papadogeorgou et al., 2023)

In this study, we demonstrate how SEM can be extended to estimate nonstationary parameters, nonlinear relationships, and statistical interactions. We specifically combine DSEM (which handles missing data as well as simultaneous and lagged relationships among variables) with MSEM (which estimates nonstationarity, nonlinearity, and statistical interactions). We also introduce a novel "path-lag-slope" notation for specifying moderated DSEM ("MDSEM") available within the R-package *dsem* version 2.0.0. Finally, we demonstrate the model and software using three varied examples. We specifically address (1) *nonstationarity* using a varying-slopes model for the relationship between temperature and regional climate; (2) *nonlinearity* using a SEM implementation of species interactions using the Lotka-Volterra equations; and (3) and *statistical interactions* using a resource-consumer-predator model involving species interactions that have a dome-shaped response to temperature.

Methods

We seek to extend dynamic structural equation models to include nonstationarity, nonlinearity, and statistical interactions. To do so, we adapt moderated SEM (MSEM) for use in the Gaussian Markov random field (GMRF) that is used when estimating dynamic SEM (DSEM). We start by reviewing this GMRF implementation of DSEM.

DSEM describes the relationship among $j \in \{1,2, \ldots J\}$ variables over $t \in \{1,2, \ldots T\}$ time-intervals, assembled in a matrix $\mathbf{X}$ with dimension $T \times J$ containing values x_{tj} for each time and variable. It is then fitted using a $T \times J$ matrix $\mathbf{Y}$ containing measurements y_{tj} (where y_{tj} might also include missing values that are specified as NA). We define operator $\mathrm{vec}(\mathbf{X})$ as stacking the columns of $\mathbf{X}$ into a single vector with length TJ, and $\mathbf{x}_t$ as the J length row-vector containing x_{tj} for all variables in time t. DSEM also specifies that variables have unmodeled sources of variation which result in exogenous variation $\mathbf{E}$ with dimension $T \times J$, where $\mathrm{vec}(\mathbf{E})$ again stacks columns into a single vector, and $\boldsymbol{\epsilon}_t$ is the error in year t. DSEM then defines a vector-autoregressive process for $\mathbf{x}_t$:

$$\mathbf{x}_t = \underbrace{\mathbf{B}_0\mathbf{x}_t}_{\text{Simultaneous}} + \underbrace{\mathbf{B}_1\mathbf{x}_{t-1}}_{\text{Lag}-1} + \underbrace{\mathbf{B}_2\mathbf{x}_{t-2}}_{\text{Lag}-2} + \underbrace{\cdots}_{\text{Higher}-\text{order}} + \boldsymbol{\epsilon}_t$$

where $\mathbf{B}_0$ are simultaneous interactions among variables, $\mathbf{B}_1$ and $\mathbf{B}_2$ are lag-1 and lag-2 interactions, and the model can include any arbitrary lag up to $T-1$. This vector-autoregressive notation is then rewritten as a joint simultaneous equation model (SEM), while assembling the lag matrices $(\mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2, \ldots)$ in a joint path matrix $\mathbf{P}_{\text{joint}}$ with dimensions $TJ \times TJ$:

$$\text{vec}(\mathbf{X}) = \mathbf{P}_{\text{joint}} \text{vec}(\mathbf{X}) + \text{vec}(\mathbf{E}) \tag{1}$$

$$\text{vec}(\mathbf{E}) \sim \text{MVN}(\mathbf{0}, \mathbf{V}_{\text{joint}})$$

where $\text{vec}(\mathbf{E})$ has variance $\mathbf{V}_{\text{joint}}$ with dimension $TJ \times TJ$. Algebra then shows that:

$$\text{vec}(\mathbf{X}) \sim \text{GMRF}(\mathbf{0}, \mathbf{Q}) \tag{2}$$

$$\mathbf{Q} = \left(\mathbf{I} - \mathbf{P}_{\text{joint}}^{\text{T}}\right) \mathbf{V}_{\text{joint}}^{-1} (\mathbf{I} - \mathbf{P}_{\text{joint}})$$

where the probability density for this GMRF can be efficiently evaluated to estimate values for $\mathbf{P}_{\text{joint}}$.

To allow users to specify what combination of simultaneous and lagged effects are estimated (i.e., the non-zero elements of lag-matrices lag matrices $\mathbf{B}_0, \mathbf{B}_1, \mathbf{B}_2, \ldots$), DSEM introduces a "path-and-lag" notation. For example, for $J = 2$ variables $\mathbf{x}_t = (A_t, B_t)$ over $T = 3$ times we might specify that changing A_t by Δ causes a lagged change of $\gamma\Delta$ in B_{t+1}, represented in path-and-lag notation as $A \rightarrow B, 1, \beta$. Each one-headed arrow in path-and-lag notation then defines a path coefficient γ_k that links two variables as represented by $J \times J$ indicator matrix $\mathbf{P}_k$, and is also associated with a lag-matrix $\mathbf{L}_{g[k]}$ where $g[k]$ is the specified lag for path coefficient k. In our example with $K = 1$ path coefficient with a lag-1 effect $g[1] = 1$, this results in:

$$\mathbf{L}_{g[1]} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \tag{3}$$

and indicator matrix:

$$\mathbf{P}_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \tag{4}$$

The joint path matrix is then assembled by summing the contribution of all K path coefficients:

$$\mathbf{P}_{\text{joint}} = \sum_{k=1}^{K} \gamma_k (\mathbf{L}_{g[k]} \otimes \mathbf{P}_k) \tag{5}$$

where $\mathbf{P}$ is a $K \times J \times J$ array containing all indicator matrices $\mathbf{P}_k$, and $\mathbf{L}_{g[k]} \otimes \mathbf{P}_k$ is the Kronecker product that results in a $TJ \times TJ$ matrix to match the size of $\mathbf{P}_{\text{joint}}$. In our example, this results in a joint-path matrix:

$$\mathbf{P}_{\text{joint}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{6}$$

where parameter γ_1 is now duplicated across times.

Moderated variables

Moderated path analysis (Klein & Moosbrugger, 2000) extends the simultaneous equation to include quadratic terms, $\mathbf{X} = \mathbf{P}_{\text{joint}}\text{vec}(\mathbf{X}) + \text{vec}(\mathbf{X})^{\text{T}}\mathbf{N}_{\text{joint}}\text{vec}(\mathbf{X})$, where $\mathbf{N}$ is an asymmetric $TJ \times TJ$ matrix that contains quadratic effects. This then allows quadratic relationships among variables, e.g., $B = \gamma A + \nu A^2$, where γ is the linear in matrix $\mathbf{P}$ and ν is the quadratic effect of A on B in $\mathbf{N}$. We specify a moderated structural equation model that has a similar property but using methods that can be written as a GMRF, as we next show. In particular, we introduce "path-lag-slope" notation, where e.g., $A \rightarrow B, 0, C$ allows a variable C to replace a stationary parameter when representing the slope linking A to B.

Specifically, we replace a stationary parameter γ_k with a specified model variable $\mathbf{x}_j$ that can then vary over time. This can be written as:

$$\mathbf{P}_{\text{joint}} = \sum_{k=1}^{K} \left(\mathbf{L}_{g[k]} \mathbf{Z}_k \right) \otimes \mathbf{P}_k \tag{7}$$

where:

$$\mathbf{Z}_k = \begin{cases} \gamma_k \mathbf{I} & \text{if path is stationary with value } \gamma_k \\ \text{diag}\left(\mathbf{x}_{j[k]}\right) & \text{if path is nonstationary with value } \mathbf{x}_{j[k]} \end{cases} \tag{8}$$

and where $\text{diag}\left(\mathbf{x}_{j[k]}\right)$ constructs a $T \times T$ diagonal matrix from vector $\mathbf{x}_{j[k]}$ for the variable $j[k]$ that replaces path coefficient k, such that $\mathbf{L}_{g[k]}\mathbf{Z}_k$ then applies a lag operator $\mathbf{L}_{g[k]}$ such that $x_{t_1 j[k]}$ is the impact of $x_{t_1 j_1}$ on $x_{t_2 j_2}$ where $t_2 \geq t_1$. We propose two interpretations for this expanded specification for DSEM (see Fig. 1 for examples):

1. *Random slopes:* The variable $\mathbf{x}_{j[k]}$ that is used instead of path coefficient k could be interpreted as a random slope (Gelman & Hill, 2007). For example, we might have a model with $J = 3$ variables, $\mathbf{x}_t = (A_t, B_t, \beta_t)$, where β_t is effect of A_t on B_t, $B_t = \mu_B + \beta_t A_t + \epsilon_t$, and written as $A \rightarrow B, 0, beta$ in arrow-lag-slope notation (see Fig. 1B).
2. *Product of two variables in a graphical model*: The joint path matrix $\mathbf{P}_{\text{joint}}$ defines a graphical model, where many edges (arrows representing causal effects) can collide in a single node (vertex representing a response variable), and this collision corresponds to adding together the contribution of these multiple predictor-variables to predict a single response-variable. However, using "moderator" variable $\mathbf{x}_{j[k]}$ to represent the impact of predictor $\mathbf{x}_{k_1}$ on response $\mathbf{x}_{k_2}$ defines a new operator in the graph, where we instead take the product of two variables $\mathbf{x}_{k_1}$ and $\mathbf{x}_{j[k]}$ when estimating their impact on $\mathbf{x}_{k_2}$. This interaction can be viewed as a "composite variable", which we visualize as a diamond to distinguish it from measured ("manifest") variables as squares or unobserved ("latent") variables as circles (see Fig. 1C). For example, we might specify $C_t = \beta_A A_t + \beta_{AB} A_t B_t + \epsilon_t$ by specifying four

variables $\mathbf{x}_t = (A_t, B_t, AB_t, C_t)$, and using arrow-lag-slope notation to specify $A \rightarrow AB, 0, B$. The resulting model can be viewed as a time-series version of a multi-level probabilistic graphical model (Koller & Friedman, 2009).

These different interpretations arise even when one or more variables has missing values.

For example, we might want to estimate the impact of temperature T on other variables. However, species densities might initially increase and then eventually decrease with increases in temperature (a "dome-shaped temperature response"). This can be approximated as a quadratic effect $Y = \beta_T T + \beta_{T^2} T^2$, and we can construct the temperature-squared term as the product of temperature and itself ($T \rightarrow T, 0, T$ in path-lag-slope notation). T^2 is then a latent variable, which is calculated from estimated (imputed) values of T. Alternatively, we can use a third-order Taylor-series (Maclaurin) approximation:

$$e^A \approx \sum_{i=0}^{3} \frac{A^i}{i!} \tag{9}$$

and using a power-series for covariate A to approximate an exponential function (Fig. 1D), with error bounded by $e^{|A|} \frac{A^4}{4!}$.

Defining the path matrix $\mathbf{P}_{\mathrm{joint}}$ using a moderated dynamic structural equation model (MDSEM) allows us to approximate interactions, exponential, and polynomial effects while also interpolating missing data, estimating direct and indirect effects, and constraining model covariance using domain knowledge. The MDSEM is available in the R package *dsem* (Thorson et al., 2024) version 2.0.0, which uses package *TMB* (Kristensen et al., 2016) to calculate the Laplace approximation (Skaug & Fournier, 2006) when calculating the marginal likelihood of parameters while integrating across random effects, and the Eigen package to efficiently calculate sparse-matrix computations (Guennebaud et al., 2010). We then optimize the marginal likelihood in the R statistical environment (R Core Team, 2023), and use a generalization of the delta method (Kass & Steffey, 1989) to calculate standard errors for parameters and derived quantities.

Simulation experiment: Autocorrelated random-slopes model

We first confirm that MDSEM can accurately recover a model variable $\mathbf{x}_{j[k]}$ that is treated as a varying slope representing the impact of one variable on another. To do so, we simulation data a "random slopes" time-series model:

$$A_t \sim Normal(0, \sigma_A^2) \tag{10}$$

$$B_t \sim Normal(\beta_t A_t, \sigma_B^2)$$

$$\beta_t = 0.5\left(\sin\left(2\pi\frac{t-1}{T-1}\right)+1\right)$$

where $\mathbf{x}_t = (A_t, B_t, \beta_t)$ is the set of $J = 3$ variables over $T = 51$ times, and slope β_t fluctuates $\beta_1 = 1$ to $\beta_{25} = 0$ and back to $\beta_{51} = 1$. We then fit MDSEM observing $\mathbf{y}_t = (A_t, B_t, \text{NA})$ and specifying a first-order autoregressive process on the slope variable. We contrast this with a conventional DSEM fitting the $J = 2$ observed variables and assuming that slope β is stationary over time. We specify $\sigma_A = 1$ and explore three scenarios that have different magnitudes of error in the response, $\sigma_B = \{0.2, 0.5, 1.0\}$. For each scenario, we simulate 500 simulation replicates, and record the estimated $\hat{\beta}_t$ (for the MDSEM) or $\hat{\beta}$ for the (DSEM), and compare these with the known true value.

Case 1: Random slopes linking regional and local habitat

We next demonstrate using MDSEM as a varying slope model using real-world data. To do so, we analyze the relationship between sea surface temperatures at Departure Bay (Vancouver Island, British Columbia, Canada) and a regional climate index (the Pacific Decadal Oscillation, PDO), using annual measurements in January from 1914-2017 (see Table S1 for code). We seek to estimate how the relationship has changed during 100 years of climate change:

$$X_t \sim \begin{cases} \text{Normal}(\mu_X, \sigma_X^2) & \text{if } t = 1914 \\ \text{Normal}(\rho_X (X_{t-1} - \mu_X) + \mu_X, \sigma_X^2) & \text{if } t > 1914 \end{cases} \tag{11}$$

$$\beta_t \sim \begin{cases} \text{Normal}(\mu_\beta, \sigma_\beta^2) & \text{if } t = 1914 \\ \text{Normal}(\rho_\beta \left(\beta_{t-1} - \mu_\beta\right) + \mu_\beta, \sigma_\beta^2) & \text{if } t > 1914 \end{cases}$$

$$Y_t \sim \begin{cases} \text{Normal}(\mu_Y, \sigma_Y^2) & \text{if } t = 1914 \\ \text{Normal}(\rho_Y (Y_{t-1} - \mu_Y) + \beta_t (X_t - \mu_X) + \mu_Y, \sigma_Y^2) & \text{if } t > 1914 \end{cases}$$

where we estimate the conditional variance and first-order autocorrelation for each of $J = 3$ variables, where PDO X_t and local temperature Y_t are both observed and β_t is a latent variable representing the time-varying slope.

Case 2: Lotka-Volterra predator-prey dynamics

We next demonstrate using MDSEM to approximate a mechanistic model that involves a nonlinear relationship among variables. To do so, we demonstrate it using the Lotka-Volterra model, which remains one of the most widely-taught descriptions for predator-prey dynamics. It defines an ordinary differential equation for the abundance of prey X_t and predators Y_t:

$$\frac{d}{dt} X_t = \alpha X_t - \beta X_t Y_t \tag{12}$$

$$\frac{d}{dt} Y_t = \gamma X_t Y_t - \delta Y_t$$

where α is the per-capita growth rate for prey, β is the prey mortality per predator-prey encounter, γ is the predator growth rate per encounter, and δ is the predator mortality rate in the absence of encounters. We first reformulate in terms of log-abundance:

$$\frac{d}{dt} \log_e(X_t) = \frac{1}{X_t} \frac{d}{dt} X_t = \alpha - \beta Y_t \tag{13}$$

$$\frac{d}{dt}\log_e(Y_t) = \frac{1}{Y_t}\frac{d}{dt}Y_t = \gamma X_t - \delta$$

and then use the 3rd-order Taylor series approximation to the exponential function, $\tilde{X}_t = \sum_{i=0}^{3} \frac{\log(X_t)^i}{i!}$ and $\tilde{Y}_t = \sum_{i=0}^{3} \frac{\log(Y_t)^i}{i!}$. Finally, we use a first-order forwards-Euler approximation to the ODE, and add process errors representing unmeasured variation in productivity for the prey ($\epsilon_{t,1}$) or predator ($\epsilon_{t,2}$):

$$\log(X_{t+1}) = \log(X_t) + \alpha - \beta\tilde{Y}_t + \epsilon_{t,1} \quad (14)$$

$$\log(Y_{t+1}) = \log(Y_t) + \gamma\tilde{X}_t - \delta + \epsilon_{t,2}$$

where this approximation can be fitted using MDSEM. We then compare MDSEM estimates $\hat{\theta}_{\mathrm{DSEM}} = (\hat{\alpha}_{\mathrm{DSEM}}, \hat{\beta}_{\mathrm{DSEM}}, \hat{\delta}_{\mathrm{DSEM}}, \hat{\gamma}_{\mathrm{DSEM}})$ with the maximum-likelihood estimate $\hat{\theta}_{\mathrm{ODE}}$ resulting from a 3rd-order Runge-Kutta ODE solver implemented using RTMB (Kristensen, 2024).

We specifically compare $\hat{\theta}_{\mathrm{DSEM}}$ and $\hat{\theta}_{\mathrm{ODE}}$ using two examples:

1. *Hare-Lynx in pelt records from Hudson Bay*: We use records of pelts for Canada Lynx and their prey snowshoe hare from Hudson Bay 1900-1920, extracted from Gotelli (2008 Fig. 6.16) and originating elsewhere (Elton & Nicholson, 1942; MacLulich, 1937);
2. *Didinium-Paramesium microcosm experiment*: We use records of *Paramesium aurelia* and *Didinium nasutum* in a microcosm experiment at 0.5 Cerophyll concentration measured every 12 hours over 35 days, i.e., $T = 71$ (Veilleux, 1979 Fig. 11a), as previously digitized (Jost & Ellner, 2000 Fig. 1).

In each case, we randomly drop 10% of measurements to demonstrate the ability to impute missing values jointly with estimating parameters.

Case 3: Temperature-dependent resource-consumer-predator dynamics

Finally, we demonstrate using MDSEM to estimate how covariates can moderate variation in slopes over time including polynomial effects. To do so, we estimate a quadratic impact of temperature on species interactions, using monthly measurements of Temperature (W_t in Celcius), *Cryptomonas* (resource, C_t in log-abundance), *Daphnia* (consumer, D_t in log-abundance), and *Leptodora* (predator, L_t in log-abundance) in Lake Washington from 1962-1994, $T = 396$ (Hampton et al., 2006). We specifically focus on dynamics for *Daphnia*, and estimate a quadratic impact of temperature on *Cryptomonas* and *Daphnia* abundance:

$$W_t = \rho_W W_{t-1} + \epsilon_{\mathrm{W},t} \tag{15}$$

$$C_t = \mu_C + \rho_C(C_{t-1} - \mu_C) + \alpha_C W_t + \beta_C W_t^2 + \epsilon_{\mathrm{C},t}$$

$$D_t = \mu_D + \rho_D(D_{t-1} - \mu_D) + \alpha_D W_t + \beta_D W_t^2 + \gamma_t(C_t - \mu_C) + \delta_t(L_{t-1} - \mu_L) + \epsilon_{\mathrm{D},t}$$

$$L_t = \mu_L + \rho_L(L_{t-1} - \mu_L) + \epsilon_{\mathrm{C},t}$$

We also estimate a simultaneous impact of *Cryptomonas* on *Daphnia* that varies over time following a quadratic temperature response:

$$\gamma_t = \mu_\gamma + \rho_\gamma(\gamma_{t-1} - \mu_\gamma) + \alpha_\gamma W_t + \beta_\gamma W_t^2 + \epsilon_{\gamma,t} \tag{16}$$

and a one-month lagged impact of *Leptodora* on *Daphnia* that also varies over time following a quadratic temperatures response:

$$\delta_t = \mu_\delta + \rho_\delta(\delta_{t-1} - \mu_\delta) + \alpha_\delta W_t + \beta_\delta W_t^2 + \epsilon_{\delta,t} \tag{16}$$

We then use a two-sided Wald test to identify which of the eight temperature parameters are statistically significant ($p < 0.05$).

Results

Simulation experiment: Autocorrelated random-slopes model

The simulation experiment confirms that MDSEM can accurately estimate autocorrelated variation for a slope parameter measuring the impact of one variable on another (Fig. 2).

However, as error in the response variable increases (from left to right panel of Fig. 2): (1) the random-slope estimate shrinks towards the average slope value across time (i.e., red line shrinks towards blue line), and (2) the simulation interval (shaded area) increases in width. Therefore, the ability of MDSEM to estimate a path coefficient as a latent variable depends upon the quality of available data.

Case 1: Random slopes linking regional and local habitat

We first confirm that a time-varying slope can be estimated using a latent variable (e.g., following an autoregressive process) as the slope parameter. In the random-slopes model predicting sea surface temperature at Departure Bay from the Pacific Decadal Oscillation (Fig. 2), the model estimates stronger autocorrelation for the varying slope (0.87) than PDO (0.49) or the conditional errors in temperature (0.11). Inspecting the estimated slope, we see the weakest association from 1915-1925, and a relatively stable slope from 1940-2017 (Fig. 3 3rd row). As expected, estimating a stationary slope (0.37) is nearly the midpoint of the estimated values when allowing the slope to be nonstationary.

Case 2: Lotka-Volterra predator-prey dynamics

We next confirm that we can use latent-moderated interactions to approximate a nonlinear (e.g., exponential) function with the widely used Lotka-Volterra model for predator-prey dynamics. Comparing interaction estimates from the MDSEM with a state-space solution to the ODE (Fig. 4), we see that the two largely agree in sign and magnitude. Differences become more pronounced for larger (> 0.5) estimated interactions, but the estimated carrying capacity is close (and within confidence intervals) for both implementations. Similarly, both models interpolate missing values similarly (Fig. 5), and in a manner that is consistent with the oscillatory dynamics of the system.

Case 3: Temperature-dependent resource-consumer-predator dynamics

Finally, we confirm that we can use latent-moderated interactions to include polynomial covariate effects, specifically specifying a quadratic temperature response on both intercepts (average density) and slopes (interactions) in a resource-consumer-predator model. Inspecting the resulting graph (Fig. 6), we see that linear and quadratic temperature effects are significant for the consumer (*Daphnia*) density, as well as the impact of the resource (*Cryptomonas*) on the consumer, whereas the other temperature responses are not statistically significant. Examining the estimated temperature-response curve (Fig. 7), the two significant effects both have a positive and dome-shaped response, where densities and interactions are highest at 12-14 degrees Celcius (Fig. 7B and 7C). By contrast, the other two temperature-response curves (Fig. 7A and 7C) have a confidence interval that could include a constant value over a large portion of the range of temperatures.

Discussion

In this paper, we demonstrated how a moderated dynamic structural equation model (MDSEM) can extend causal analysis to include nonstationarity, nonlinearity, and statistical interactions, while also interpolating missing values, specifying latent variables, and estimating both simultaneous and lagged relationships among variables. We used a simulation experiment to confirm that sinusoidal variation in a slope linking two variables can be estimated as a autoregressive latent variable, and that the estimated slope is shrunk towards a constant value as data become less informative. We then demonstrated nonstationarity using a time-varying relationship between local and regional climate, nonlinearity using Lotka-Volterra dynamics, and statistical interactions by estimating temperature-dependent interactions in a resource-consumer-predator system. The method is available in an R-package *dsem* (starting with release 2.0.0), and we next discuss how this MDSEM might be useful for a range of theoretical and applied questions throughout ecology.

SEM (and its spatial, phylogenetic, and time-series extensions) are useful for applied ecologists because they address several drawbacks of conventional linear regression (e.g., generalized linear models and analysis of variance). In particular, SEM addresses the problem of collinearity (Dormann et al., 2013) by using domain knowledge to inform predictions given novel combinations of predictors, and also accounts for missing values for both predictor and response variables by specifying a joint distribution for both. Despite these advantages, phylogenetic, spatial, and time-series applications of SEM have previously lacked any capability to estimate nonlinearity, nonstationarity, and statistical interactions, which are also of general interest in ecology. By addressing these challenges, moderated SEM seems suitable for the wide range of uses discussed in recent ecological reviews (Arif & MacNeil, 2022; Byrnes & Dee, 2025; Grace, 2024; Larsen et al., 2019). In particular, DSEM has previously required assuming that path coefficients (i.e., simultaneous and lagged interactions among variables) are constant over time. After developing a DSEM based on scientific knowledge, we recommend that analysts sequentially test the model when replacing each path coefficient with a model variable (e.g., which follows an autoregressive process), and use model selection to evaluate the strength of evidence that the path coefficient is stationary or varies over time.

In addition to a growing interest in causal analysis using SEM and PA, ecologists use custom-built hierarchical models (i.e., integrated-population or stock-assessment models [IPMs]) to predict the likely effect of hypothetical policy changes (Kéry & Schaub, 2021). IPMs are a powerful tool for applied ecologists because they allow analysts to incorporate nonlinear and state-space features that are specifically suited to their study system. However, IPMs are often specified using Bayesian hierarchical modelling, which then requires specifying a directed acyclic graph (DAG) for linkages among system components (i.e., avoiding cyclic dependencies among system variables). By contrast, SEM (including DSEM and MDSEM) can estimate these cyclic dependencies ($A \rightarrow B \rightarrow C \rightarrow A$) while simultaneously imputing missing variables. Additionally, the "arrow-lag-slope" notation developed here continues to provide a high-level interface for specifying system linkages using MDSEM. We believe that a simple and expressive interface is necessary for broad adoption of any statistical tool for ecologists, similar to how the `formula` interface for linear models (Wilkinson & Rogers, 1973) has led to broad adoption among ecologists of linear mixed and generalized additive models. Ultimately, we envision embedding MDSEM as an interface to specify linkages among process errors and/or covariates for use within IPMs (Champagnat et al., 2025).

Finally, we showed how the output of MDSEM can be plotted to summarize context-dependent and nonstationary relationships (e.g., Fig 2, 6, and S1). However, analysts will also want to compute the total effect of an exogenous (policy) change in system variables. In conventional DSEM, the total effect is computed from the $\mathbf{P}_{\text{joint}}$, which is assumed to be stationary over time. Specifically, an analyst might envision a policy that changes the states $\mathbf{X}$ to $\mathbf{X} + \mathbf{D}$, where change-matrix $\mathbf{D}$ could represent a pulse experiment (i.e., non-zero values in only a single time) or press experiment (i.e., non-zero values continuing indefinitely). This change causes in a first-order effect $\mathbf{P}_{\text{joint}}\text{vec}(\mathbf{D})$, which in turn causes a second-order effect $\mathbf{P}_{\text{joint}}^2\text{vec}(\mathbf{D})^T$, and where the total effects is then $\left(\mathbf{I} - \mathbf{P}_{\text{joint}}\right)^{-1}\text{vec}(\mathbf{D})$. By contrast, MDSEM allows $\mathbf{P}_{\text{joint}}$ to vary due to other latent or endogenous variables. Computing the total effect therefore involves a first-order effect, $\mathbf{P}_1\text{vec}(\mathbf{D})$ where $\mathbf{P}_1 = \mathbf{P}_{\text{joint}}$. However, the second-order effect requires updated path matrix $\mathbf{P}_2$, calculated by updating $\mathbf{P}_{\text{joint}}$ given the previous first-order effects (Eq. 5A-5B), where the second-order effect is $\mathbf{P}_2\mathbf{P}_1\text{vec}(\mathbf{D})$. By extension, the total effect is the sum across all such partial effects, $\text{vec}(\mathbf{D})^T(\sum_{k=1}^{\infty}\prod_{k'=1}^{k}\mathbf{P}_{k'})$. We therefore acknowledge that interpreting the total effect is more complicated in moderated SEM than in conventional cases.

References

Arif, S., & MacNeil, M. A. (2022). Predictive models aren't for causal inference. *Ecology Letters*, *25*(8), 1741–1745. https://doi.org/10.1111/ele.14033

Arif, S., & MacNeil, M. A. (2023). Applying the structural causal model framework for observational causal inference in ecology. *Ecological Monographs*, *93*(1), e1554. https://doi.org/10.1002/ecm.1554

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Byrnes, J. E. K., & Dee, L. E. (2025). Causal Inference With Observational Data and Unobserved Confounding Variables. *Ecology Letters*, *28*(1), e70023. https://doi.org/10.1111/ele.70023

Champagnat, J., Monnahan, C. C., Sullivan, J. Y., Thorson, J. T., Shotwell, K. S., Rogers, L. A., & Punt, A. E. (2025). *Causal models as a scientific framework for next-generation ecosystem and climate-linked stock assessments*. https://ecoevorxiv.org/repository/view/9794/

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013).

Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Elton, C., & Nicholson, M. (1942). The Ten-Year Cycle in Numbers of the Lynx in Canada. *Journal of Animal Ecology*, *11*(2), 215–244. https://doi.org/10.2307/1358

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gotelli, N. J. (2008). *A Primer of Ecology, Fourth Edition by Nicholas J. Gotelli*. Sinauer Associates, 2008.

Grace, J. B. (2024). An integrative paradigm for building causal knowledge. *Ecological Monographs*, *94*(4), e1628. https://doi.org/10.1002/ecm.1628

Guennebaud, G., Jacob, B., & others. (2010). *Eigen v3*. http://eigen.tuxfamily.org

Hampton, S. E., Scheuerell, M. D., & Schindler, D. E. (2006). Coalescence in the Lake Washington story: Interaction strengths in a planktonic food web. *Limnology and Oceanography*, *51*(5), 2042–2051. https://doi.org/10.4319/lo.2006.51.5.2042

Hixon, M. A., & Carr, M. H. (1997). Synergistic predation, density dependence, and population regulation in marine fish. *Science*, *277*(5328), 946.

Holmes, E. E., Ward, E. J., & Wills, K. (2012). Marss: Multivariate autoregressive state-space models for analyzing time-series data. *The R Journal*, *4*, 11–19.

Ives, A. R. (2022). Random errors are neither: On the interpretation of correlated data. *Methods in Ecology and Evolution*, *13*(10), 2092–2105. https://doi.org/10.1111/2041-210X.13971

Ives, A. R., Dennis, B., Cottingham, K. L., & Carpenter, S. R. (2003). Estimating community stability and ecological interactions from time-series data. *Ecological Monographs*, *73*(2), 301–330.

Jost, C., & Ellner, S. P. (2000). Testing for predator dependence in predator-prey dynamics: A non-parametric approach. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. https://doi.org/10.1098/rspb.2000.1186

Kass, R. E., & Steffey, D. (1989). Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical bayes models). *Journal of the American Statistical Association*, *84*(407), 717–726. https://doi.org/10.2307/2289653

Kéry, M., & Schaub, M. (2021). *Integrated population models: Theory and ecological applications with R and JAGS* (1st edition). Academic Press.

Klein, A., & Moosbrugger, H. (2000). Maximum Likelihood Estimation of Latent Interaction Effects with the LMS Method. *Psychometrika*, *65*(4), 457–474. https://doi.org/10.1007/BF02296338

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. MIT press. https://books.google.com/books?hl=en&lr=&id=7dzpHCHzNQ4C&oi=fnd&pg=PR9&dq=Probabilistic+Graphical+Models+Principles+and+Techniques&ots=py4Cxn2-zM&sig=gptWuvP1MZ_ebACxp7A71XGOtzM

Kristensen, K. (2024). *RTMB: "R" Bindings for "TMB."* https://CRAN.R-project.org/package=RTMB

Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic differentiation and Laplace approximation. *Journal of Statistical Software*, *70*(5), 1–21. https://doi.org/10.18637/jss.v070.i05

Larsen, A. E., Meng, K., & Kendall, B. E. (2019). Causal analysis in control–impact ecological studies with observational data. *Methods in Ecology and Evolution*, *10*(7), 924–934. https://doi.org/10.1111/2041-210X.13190

Lefcheck, J. S. (2016). piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573–579. https://doi.org/10.1111/2041-210X.12512

MacLulich, D. A. (1937). *Fluctuations in the Numbers of the Varying Hare (Lepus Americanus)*. University of Toronto Press. https://doi.org/10.3138/9781487583064

Mahaffy, J. (2009). *Qualitative Analysis of the Lotka–Volterra Equations*. San Diego State University. https://jmahaffy.sdsu.edu/courses/f09/math636/lectures/lotka/qualde2.html

Munch, S. B., Giron-Nava, A., & Sugihara, G. (2018). Nonlinear dynamics and noise in fisheries recruitment: A global meta-analysis. *Fish and Fisheries*, *19*(6), 964–973. https://doi.org/10.1111/faf.12304

Papadogeorgou, G., Bello, C., Ovaskainen, O., & Dunson, D. B. (2023). Covariate-Informed Latent Interaction Models: Addressing Geographic & Taxonomic Bias in Predicting Bird–Plant Interactions. *Journal of the American Statistical Association*, *118*(544), 2250–2261. https://doi.org/10.1080/01621459.2023.2208390

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rollinson, C. R., Finley, A. O., Alexander, M. R., Banerjee, S., Dixon Hamil, K.-A., Koenig, L. E., Locke, D. H., DeMarche, M. L., Tingley, M. W., Wheeler, K., Youngflesh, C., & Zipkin, E. F. (2021). Working across space and time: Nonstationarity in ecological research and application. *Frontiers in Ecology and the Environment*, *19*(1), 66–72. https://doi.org/10.1002/fee.2298

Skaug, H., & Fournier, D. (2006). Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. *Computational Statistics & Data Analysis*, *51*(2), 699–709.

Thorson, J. T., Anderson, S. C., Goddard, P., & Rooper, C. N. (2025). tinyVAST: R Package With an Expressive Interface to Specify Lagged and Simultaneous Effects in Multivariate Spatio-Temporal Models. *Global Ecology and Biogeography*, *34*(4), e70035. https://doi.org/10.1111/geb.70035

Thorson, J. T., Andrews III, A. G., Essington, T. E., & Large, S. I. (2024). Dynamic structural equation models synthesize ecosystem dynamics constrained by ecological mechanisms. *Methods in Ecology and Evolution*, *15*(4), 744–755. https://doi.org/10.1111/2041-210X.14289

Thorson, J. T., Ciannelli, L., & Litzow, M. A. (2020). Defining indices of ecosystem variability using biological samples of fish communities: A generalization of empirical orthogonal functions. *Progress in Oceanography*, *181*, 102244. https://doi.org/10.1016/j.pocean.2019.102244

Thorson, J. T., Maureaud, A. A., Frelat, R., Mérigot, B., Bigman, J. S., Friedman, S. T., Palomares, M. L. D., Pinsky, M. L., Price, S. A., & Wainwright, P. (2023). Identifying direct and indirect associations among traits by merging phylogenetic comparative methods and structural equation models. *Methods in Ecology and Evolution*, *14*(5), 1259–1275. https://doi.org/10.1111/2041-210X.14076

Veilleux, B. G. (1979). An Analysis of the Predatory Interaction Between Paramecium and Didinium. *Journal of Animal Ecology*, *48*(3), 787–803. https://doi.org/10.2307/4195

von Hardenberg, A., & Gonzalez-Voyer, A. (2013). Disentangling evolutionary cause-effect relationships with phylogenetic confirmatory path analysis. *Evolution; International Journal of Organic Evolution*, *67*(2), 378–387. https://doi.org/10.1111/j.1558-5646.2012.01790.x

Wilkinson, G. N., & Rogers, C. E. (1973). Symbolic Description of Factorial Models for Analysis of Variance. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *22*(3), 392–399. https://doi.org/10.2307/2346786

Wootton, J. T., & Emmerson, M. (2005). Measurement of interaction strength in nature. *Annual Review of Ecology, Evolution, and Systematics*, 419–444.

Fig. 1 – Graphical models illustrating potential uses for latent-moderated dynamic structural equation models (MDSEM), where measured (“manifest”) variables are boxes, unmeasured (“latent”) variables are circles, varying slopes (“latent-moderated paths”) are diamonds, and arrows point from predictor to response variable while listing either a Greek symbol (representing an estimated parameter) or a Arabic numeral (representing a value that is fixed a priori), and also showing the resulting equations below each panel. We contrast the simple case of a regression model with two independent predictors (panel A), a regression with a statistical interaction (panel B), a regression with a randomly varying slope (panel C), and a 3rd-order Taylor series approximation to the exponential function (panel D).

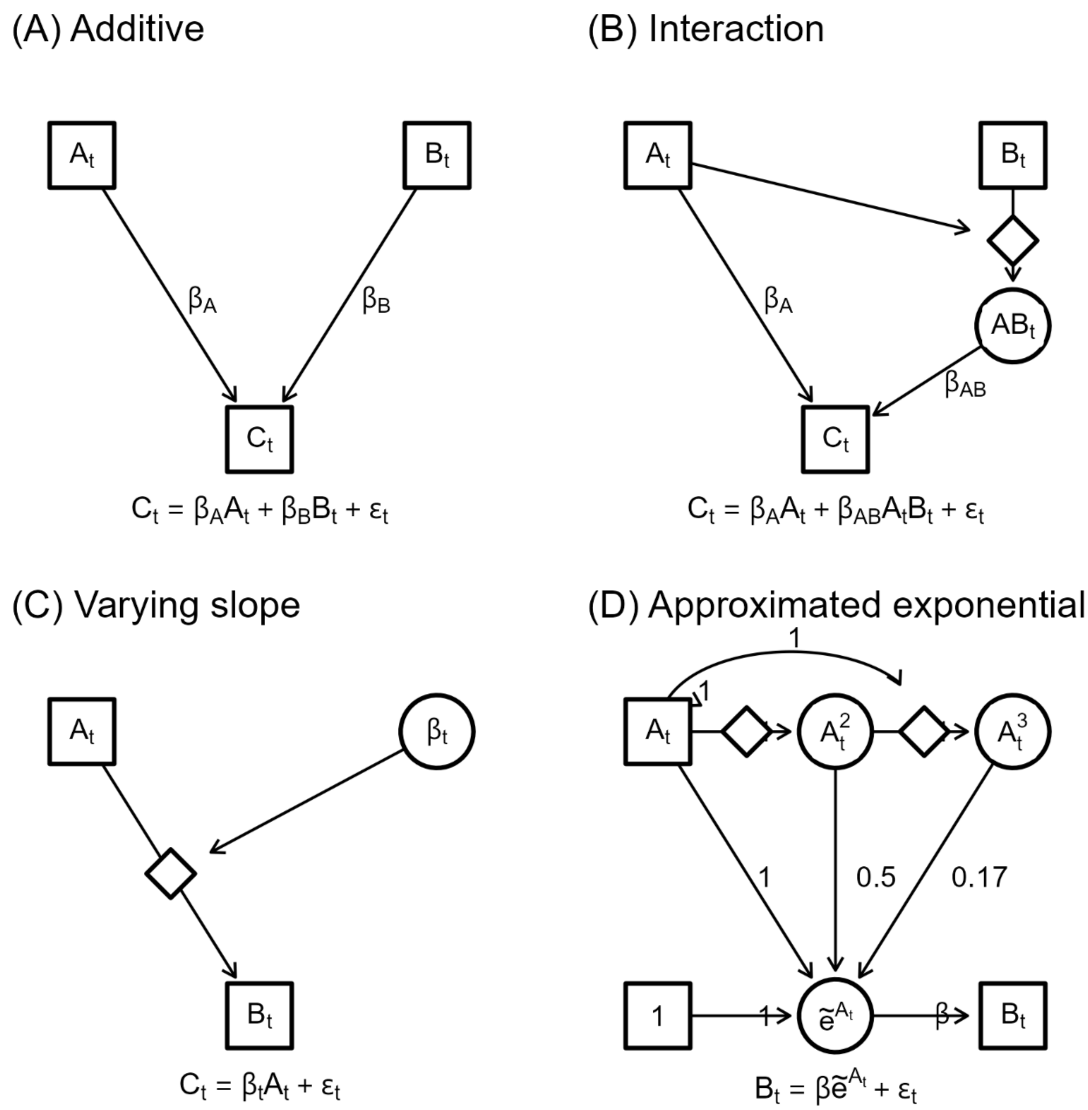

Fig. 2 – Results from a simulation experiment involving a random slope model, $A_t = \beta_t B_t + \epsilon_t$ where slope β_t (y-axis) follows a sinusoidal pattern (black line) over 50 times (x-axis), and we vary the standard deviation of process errors ϵ_t from low ($\sigma_B = 0.2$, left panel) to medium ($\sigma_B = 0.5$, middle panel) or high ($\sigma_B = 1$, right panel) levels, while estimating either a first-order autoregressive process for the slope (red line and shading) or a constant slope (blue line and shading), where the lines show the mean across 500 simulation replicates, and the shading shows the 10% and 90% simulation interval for each model.

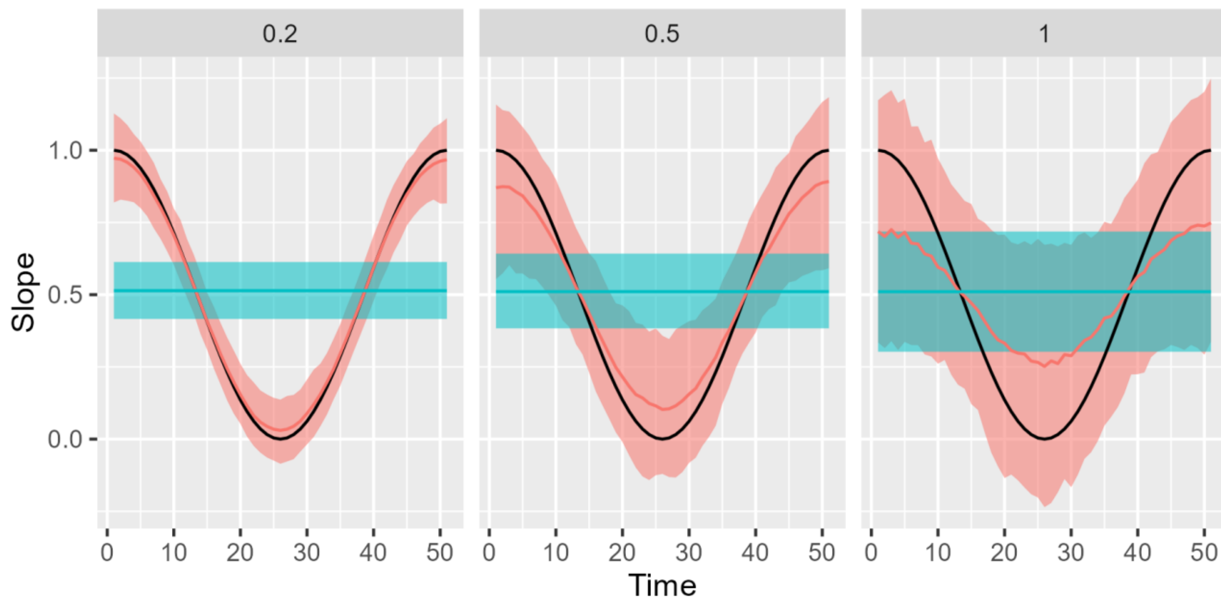

Fig. 2 – Graphical model and parameter estimates for case study #1, where the Pacific Decadal Oscillation (P_t) is used to predict sea surface temperature at Departure Bay (T_t) with a slope (β_t) that varies as a first-order autoregressive process over time (see Fig. 1 for details about graphical notation).

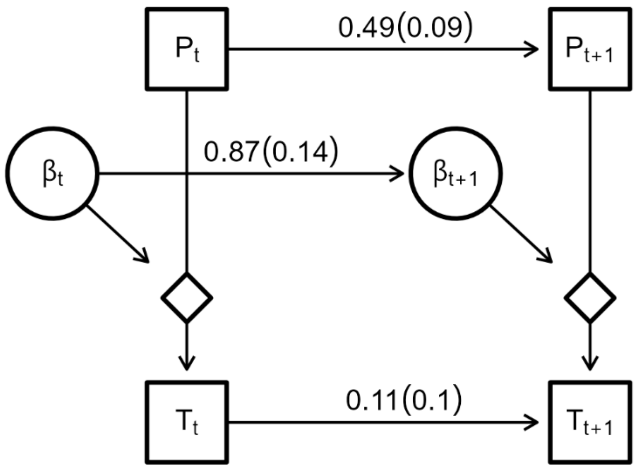

Fig. 3 – Observed values (circles), estimated values (lines), and 95% confidence intervals (shaded area) for case study #1, showing the Pacific Decadal Oscillation (PDO; top panel), Temperature (2nd panel), and estimated slope (3rd panel), contrasted with the slope estimated by an alternative model when assuming that it is stationary over time (4th panel), and using a shared y-axis scale for the slope estimates (3rd and 4th panels).

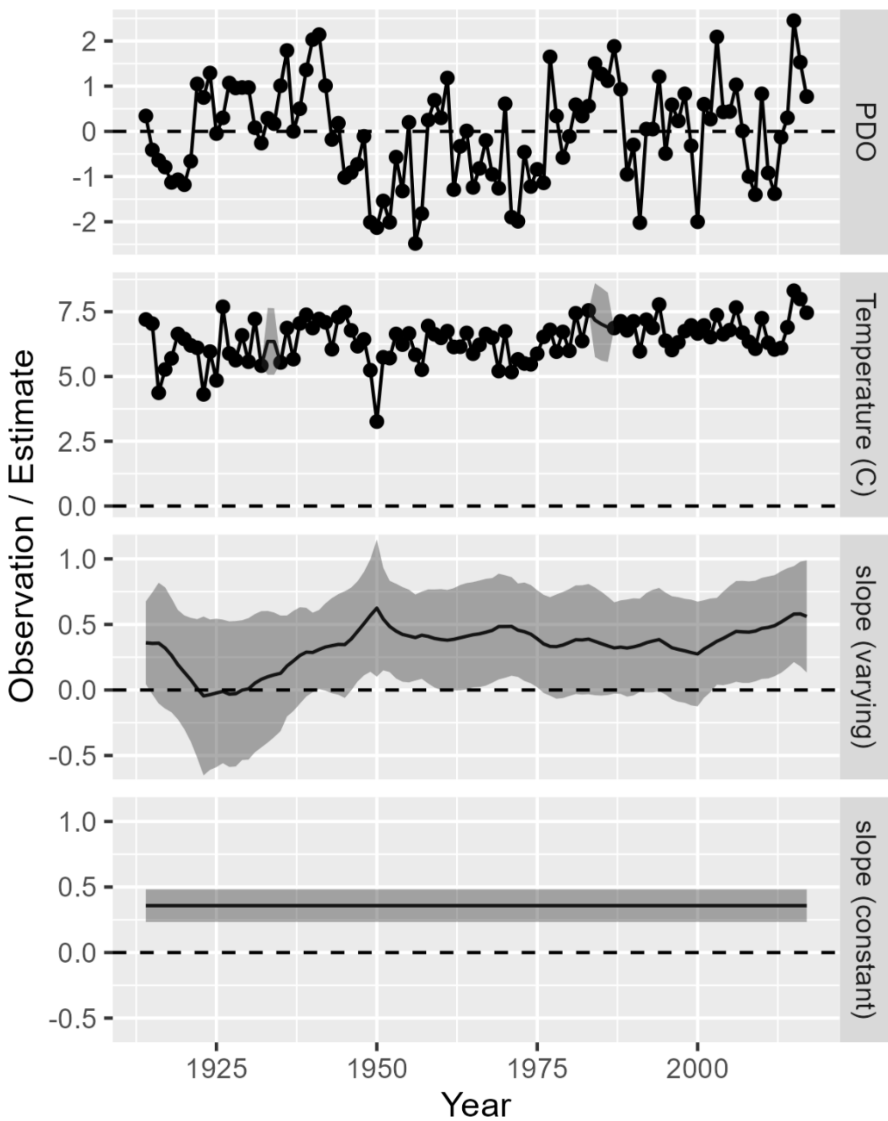

Fig. 4 – Estimated interaction parameters ($\alpha, \beta, \gamma, \delta$, dots) and 95% confidence intervals (whiskers) for case study #2 involving Lotka-Volterra dynamics, as well as the predicted carrying capacity for the prey $K_x = \frac{\alpha}{beta}$ and predator $K_y = \frac{\gamma}{\delta}$, estimated using the latent-moderated dynamic structural equation model with a Taylor-series approximation to a nonlinear (exponential) function (red) or a state-space ODE model (blue), for each of two case studies involving Hare-Lynx dynamics in Hudson Bay (top panel), or a Paramesium-Didinium microcosm experiment (bottom panel)

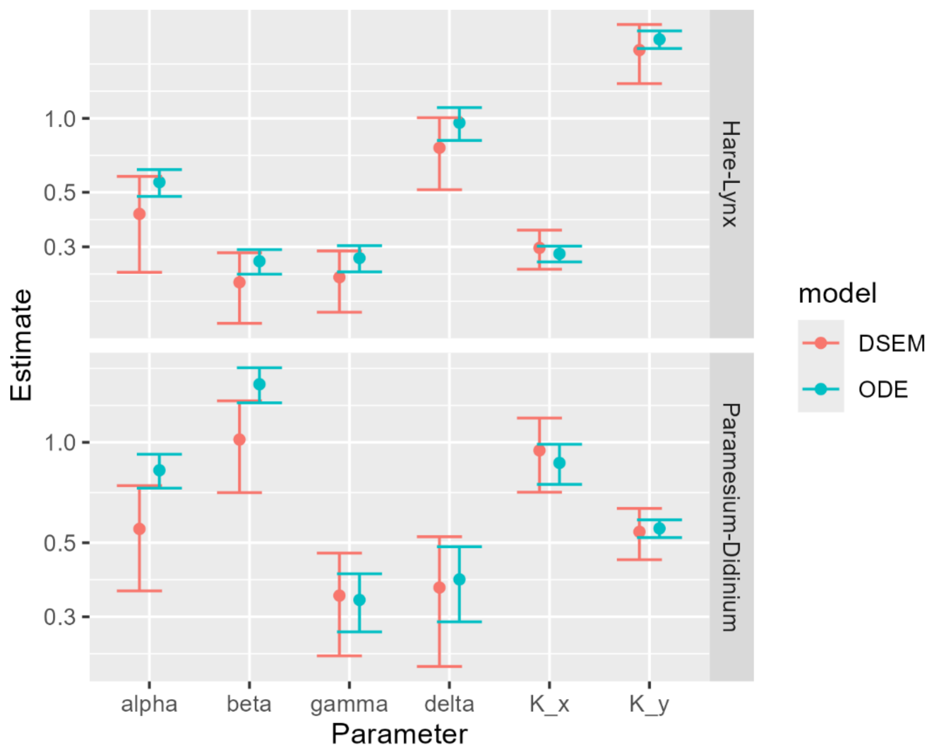

Fig. 5 -- Observed values (circles), estimated values (lines), and 95% confidence intervals (shaded area) for case study #2 involving Lotka-Volterra dynamics, contrasting estimates using the latent-moderated dynamic structural equation model (red) or the state-space ODE model (blue). Note that the MDSEM assumes that measurements are provided without error and hence only shows confidence intervals for the 10% of observations that were randomly selected and dropped prior to fitting the model.

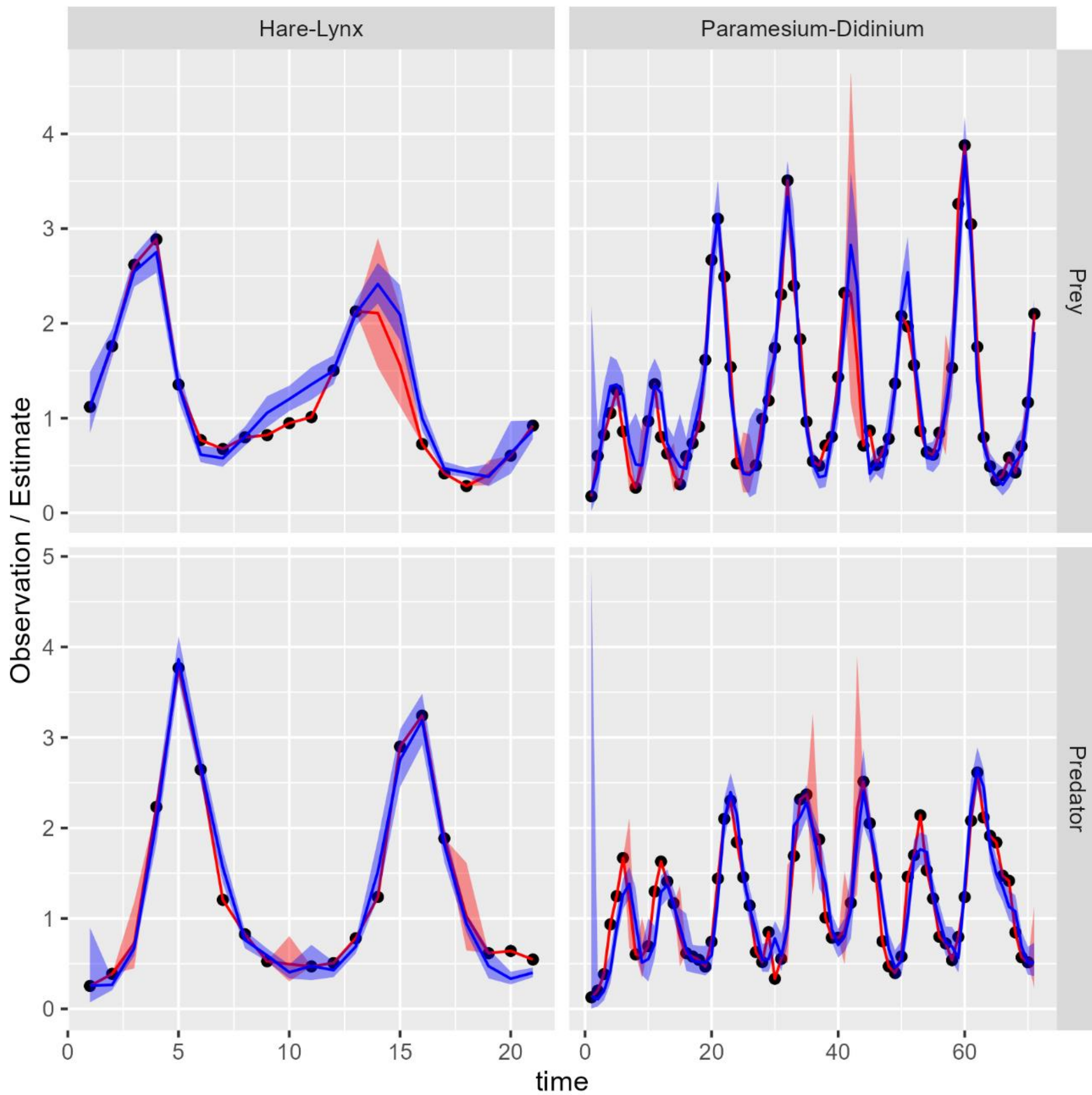

Fig. 6 – Graphical model and parameter estimates for case study #3 involving temperature-dependent resource-consumer-predator interactions (see Fig. 1 for details about graphical notation), showing Temperature T_t and its polynomial expansion as latent variable T_t^2, resource Chryptomonas C_t, consumer *Dapnia* D_t, and predator *Leptodora* L_t, and showing the time-varying impact of resource on consumers γ_t or predators on consumers δ_t. We also distinguish linkages (arrows) that are statistically significant (black arrows) or not (grey arrows).

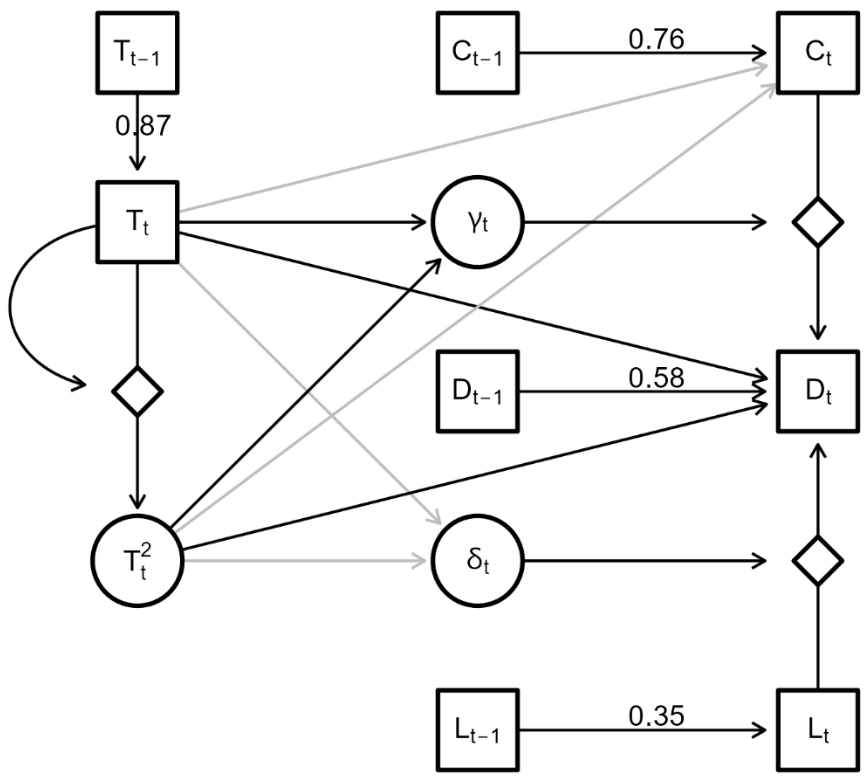

Fig. 7 – Estimated quadratic temperature-response curves (lines) and 95% confidence intervals (shaded area), showing the temperature impact on resource (top-left), consumers (top-right), the impact of resource on consumers (bottom-left), and the impact of predators on consumers (bottom-right)

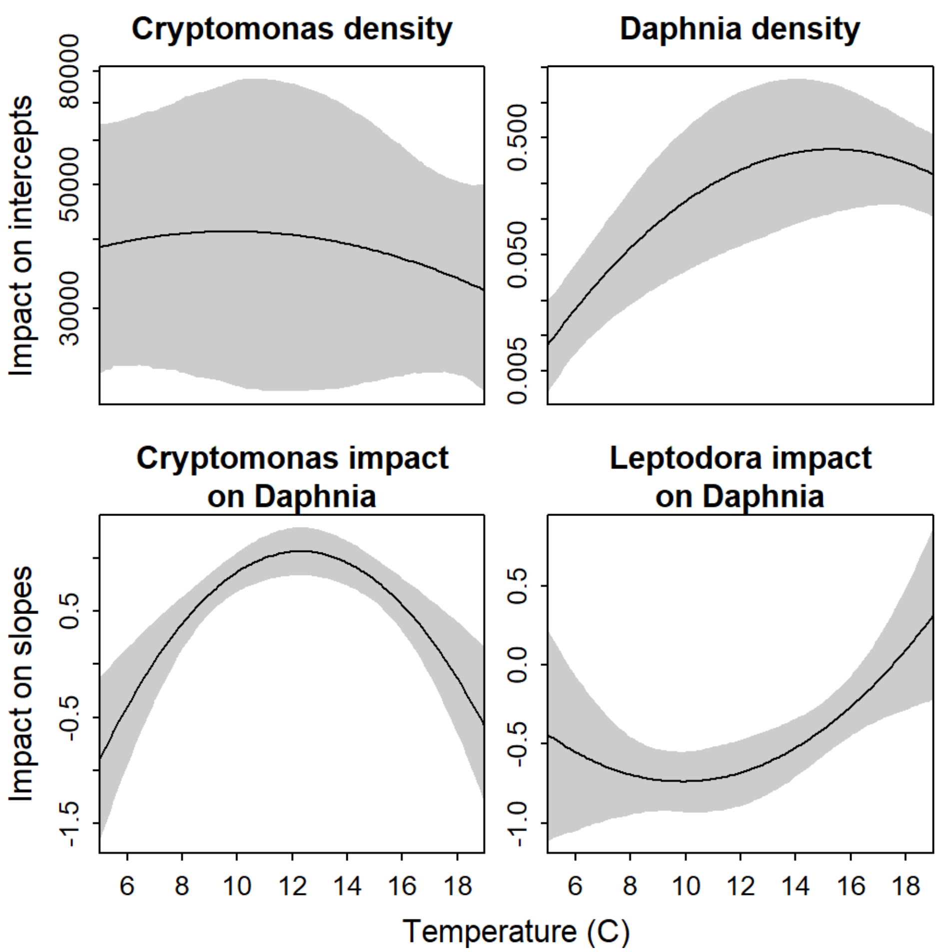

Table S1: Illustrating the code used to fit case study #1, including the arrow-lag-slope notation (left column), defining a time-series object including all variables (right column top), and the call to package *dsem* to fit the model (right column bottom).

Case 1: Varying slopes

```
# Model
sem = "
  PDO -> Temp, 0, slope
  slope -> slope, 1, ar_slope
  PDO -> PDO, 1, ar_PDO
  Temp -> Temp, 1, ar_Temp
"
```

```
# Data
tsdata = ts(data.frame(
  Temp = Combo[,2],
  PDO = Combo[,3],
  slope = NA
), start = 1914 )

# Fit
fit = dsem(
  tsdata = tsdata,
  sem = sem,
  estimate_mu = c("Temp","PDO","slope")
)
```

Table S2: Illustrating the code used to fit case study #2 (see Table S1 caption for more details).

Case 2: Lotka-Volterra

```
# Model
sem = "
  # Main interactions
  logX -> logX, 1, NA, 1
  ones -> logX, 0, alpha
  Y -> logX, 1, beta, -0.1

  # Form X \approx exp(logX)
  ones -> X, 0, NA, 1
  logX -> logX1, 0, NA, 1
  logX1 -> X, 0, NA, 1
  logX1 -> logX2, 0, logX
  logX2 -> X, 0, NA, 0.5
  logX2 -> logX3, 0, logX
  logX3 -> X, 0, NA, 0.166

  # Variances
  X <-> X, 0, NA, 0.001
  logX <-> logX, 0, sd_logX
  logX1 <-> logX1, 0, NA, 0.001
  logX2 <-> logX2, 0, NA, 0.001
  logX3 <-> logX3, 0, NA, 0.001

  # Main interactions
  logY -> logY, 1, NA, 1
  X -> logY, 1, gamma
  ones -> logY, 0, delta, -0.1

  # Form Y \approx exp(logY)
  ones -> Y, 0, NA, 1
  logY -> logY1, 0, NA, 1
  logY1 -> Y, 0, NA, 1
  logY1 -> logY2, 0, logY
  logY2 -> Y, 0, NA, 0.5
  logY2 -> logY3, 0, logY
  logY3 -> Y, 0, NA, 0.166

  # Variances
  Y <-> Y, 0, NA, 0.001
  logY <-> logY, 0, sd_logY
  logY1 <-> logY1, 0, NA, 0.001
  logY2 <-> logY2, 0, NA, 0.001
  logY3 <-> logY3, 0, NA, 0.001

# Data
Z = cbind(
  logX = log(dat$X),
  logY = log(dat$Y),
  X = NA,
  Y = NA,
  logX1 = NA,
  logY1 = NA,
  logX2 = NA,
  logY2 = NA,
  logX3 = NA,
  logY3 = NA,
  ones = 1
)

# Fit
fit = dsem(
  tsdata = ts(Z),
  sem = sem,
  estimate_mu = vector(),
)
```

```
  # Dummy constant
  ones <-> ones, 0, NA, 0.001
  ones -> ones, 1, NA, 1
"
```

Table S3: Illustrating the code used to fit case study #3 (see Table S1 caption for more details).

Case 3: Temperature-dependent resource-consumer-predator

```
# Model
sem = "
  # Temperature effect on resource density
  Temp -> Cryptomonas, 0, T_to_C
  Temp2 -> Cryptomonas, 0, T2_to_C

  # Temperature effect on consumer density
  Temp -> Daphnia, 0, T_D
  Temp2 -> Daphnia, 0, T2_D

  # Impacts on consumer
  Cryptomonas -> Daphnia, 0, alpha  #  C_D
  Leptodora -> Daphnia, 1, beta   # alpha

  # Density dependence
  Cryptomonas -> Cryptomonas, 1, ar_C
  Daphnia -> Daphnia, 1, ar_D
  Leptodora -> Leptodora, 1, ar_L

  # Form Temp^2
  Temp -> Temp2, 0, Temp
  Temp2 <-> Temp2, 0, NA, 0.001

  # Temperature on resource-consumer slope
  alpha <-> alpha, 0, NA, 0.001
  Temp -> alpha, 0, T_alpha
  Temp2 -> alpha, 0, T2_alpha

  # Temperature on predator-consumer slope
  beta <-> beta, 0, NA, 0.001
  Temp -> beta, 0, T_beta
  Temp2 -> beta, 0, T2_beta
"
```

```
# Data
Z = ts(cbind(
  dat$Temp,
  dat$Daphnia,
  dat$Leptodora,
  dat$Cryptomonas,
  alpha = NA,
  Temp2 = NA,
  beta = NA
), start = 1962, freq = 12)

# Fit
fit = dsem(
  tsdata = Z,
  sem = sem,
  estimate_mu = c(
    "Daphnia",
    "Leptodora",
    "Cryptomonas",
    "alpha",
    "beta"
  )
)
```

Table S4: Estimated path coefficients for the Departure Bay case study involving a varying-slope model linking the Pacific Decadal Oscillation (PDO) to temperatures at a lighthouse near Departure Bay, listing the model path (first column), time lag (2nd column), parameter name (3rd column), maximum-likelihood estimate and asymptotic standard error (4th and 5th columns), and the z-value and p-value from a two-sided Wald test (6th and 7th columns), where columns 4-7 are NA for parameters that are either fixed, or which vary over time (i.e., the parameter Name matches a model variable, such as the 1st row).

Path	Lag	Name	Estimate	Std_Error	z_value	p_value
PDO -> Temp	0	slope	NA	NA	NA	NA
slope -> slope	1	ar_slope	0.869	0.142	6.105	0
PDO -> PDO	1	ar_PDO	0.487	0.088	5.529	0
Temp -> Temp	1	ar_Temp	0.112	0.099	1.133	0.257
Temp <-> Temp	0	V[Temp]	0.637	0.055	11.618	0
PDO <-> PDO	0	V[PDO]	0.967	0.067	14.339	0
slope <-> slope	0	V[slope]	0.119	0.082	1.453	0.146

Table S5: Estimated path coefficients for the Lynx-Hare case study involving Lotka-Volterra predator-prey dynamics (see Table S4 caption for details)

Path	Lag	Name	Estimate	Std_Error	z_value	p_value
logX -> logX	1	NA	1	NA	NA	NA
ones -> logX	0	alpha	0.39	0.079	4.925	0
Y -> logX	1	beta	-0.32	0.05	-6.382	0
ones -> X	0	NA	1	NA	NA	NA
logX -> logX1	0	NA	1	NA	NA	NA
logX1 -> X	0	NA	1	NA	NA	NA
logX1 -> logX2	0	logX	NA	NA	NA	NA
logX2 -> X	0	NA	0.5	NA	NA	NA
logX2 -> logX3	0	logX	NA	NA	NA	NA
logX3 -> X	0	NA	0.166	NA	NA	NA
X <-> X	0	NA	0.001	NA	NA	NA
logX <-> logX	0	sd_logX	0.225	0.037	6.025	0
logX1 <-> logX1	0	NA	0.001	NA	NA	NA
logX2 <-> logX2	0	NA	0.001	NA	NA	NA
logX3 <-> logX3	0	NA	0.001	NA	NA	NA
logY -> logY	1	NA	1	NA	NA	NA
X -> logY	1	gamma	0.639	0.105	6.075	0
ones -> logY	0	delta	-0.765	0.142	-5.372	0
ones -> Y	0	NA	1	NA	NA	NA
logY -> logY1	0	NA	1	NA	NA	NA
logY1 -> Y	0	NA	1	NA	NA	NA
logY1 -> logY2	0	logY	NA	NA	NA	NA
logY2 -> Y	0	NA	0.5	NA	NA	NA
logY2 -> logY3	0	logY	NA	NA	NA	NA
logY3 -> Y	0	NA	0.166	NA	NA	NA
Y <-> Y	0	NA	0.001	NA	NA	NA
logY <-> logY	0	sd_logY	0.352	0.06	5.908	0
logY1 <-> logY1	0	NA	0.001	NA	NA	NA
logY2 <-> logY2	0	NA	0.001	NA	NA	NA
logY3 <-> logY3	0	NA	0.001	NA	NA	NA
ones <-> ones	0	NA	0.001	NA	NA	NA
ones -> ones	1	NA	1	NA	NA	NA

Table S6: Estimated path coefficients for the Didinium-Paramesium case study involving Lotka-Volterra predator-prey dynamics (see Table S4 caption for details)

Path	Lag	Name	Estimate	Std_Error	z_value	p_value
logX -> logX	1	NA	1	NA	NA	NA
ones -> logX	0	alpha	0.523	0.105	5.001	0
Y -> logX	1	beta	-0.44	0.078	-5.611	0
ones -> X	0	NA	1	NA	NA	NA
logX -> logX1	0	NA	1	NA	NA	NA
logX1 -> X	0	NA	1	NA	NA	NA
logX1 -> logX2	0	logX	NA	NA	NA	NA
logX2 -> X	0	NA	0.5	NA	NA	NA
logX2 -> logX3	0	logX	NA	NA	NA	NA
logX3 -> X	0	NA	0.166	NA	NA	NA
X <-> X	0	NA	0.001	NA	NA	NA
logX <-> logX	0	sd_logX	0.429	0.038	11.143	0
logX1 <-> logX1	0	NA	0.001	NA	NA	NA
logX2 <-> logX2	0	NA	0.001	NA	NA	NA
logX3 <-> logX3	0	NA	0.001	NA	NA	NA
logY -> logY	1	NA	1	NA	NA	NA
X -> logY	1	gamma	0.302	0.058	5.238	0
ones -> logY	0	delta	-0.361	0.083	-4.351	0
ones -> Y	0	NA	1	NA	NA	NA
logY -> logY1	0	NA	1	NA	NA	NA
logY1 -> Y	0	NA	1	NA	NA	NA
logY1 -> logY2	0	logY	NA	NA	NA	NA
logY2 -> Y	0	NA	0.5	NA	NA	NA
logY2 -> logY3	0	logY	NA	NA	NA	NA
logY3 -> Y	0	NA	0.166	NA	NA	NA
Y <-> Y	0	NA	0.001	NA	NA	NA
logY <-> logY	0	sd_logY	0.406	0.036	11.15	0
logY1 <-> logY1	0	NA	0.001	NA	NA	NA
logY2 <-> logY2	0	NA	0.001	NA	NA	NA
logY3 <-> logY3	0	NA	0.001	NA	NA	NA
ones <-> ones	0	NA	0.001	NA	NA	NA
ones -> ones	1	NA	1	NA	NA	NA

Table S7: Estimated path coefficients for the Lake Washington case study involving temperature-dependent resource-consumer-predator dynamics (see Table S4 caption for details)

Path	Lag	Name	Estimate	Std_Error	z_value	p_value
Temp -> Cryptomonas	0	T_to_C	-0.001	0.024	-0.06	0.952
Temp2 -> Cryptomonas	0	T2_to_C	-0.003	0.006	-0.498	0.618
Temp -> Daphnia	0	T_D	0.385	0.074	5.196	0
Temp2 -> Daphnia	0	T2_D	-0.036	0.015	-2.378	0.017
Leptodora -> Daphnia	1	beta	NA	NA	NA	NA
Cryptomonas -> Daphnia	0	alpha	NA	NA	NA	NA
Cryptomonas -> Cryptomonas	1	ar_C	0.758	0.038	19.716	0
Daphnia -> Daphnia	1	ar_D	0.577	0.036	16.144	0
Leptodora -> Leptodora	1	ar_L	0.348	0.077	4.533	0
Temp -> Temp2	0	Temp	NA	NA	NA	NA
Temp2 <-> Temp2	0	NA	0.001	NA	NA	NA
alpha <-> alpha	0	NA	0.001	NA	NA	NA
Temp -> alpha	0	T_alpha	0.169	0.044	3.848	0
Temp2 -> alpha	0	T2_alpha	-0.037	0.009	-4.171	0
beta <-> beta	0	NA	0.001	NA	NA	NA
Temp -> beta	0	T_beta	0.003	0.037	0.094	0.925
Temp2 -> beta	0	T2_beta	0.013	0.007	1.809	0.07
Temp <-> Temp	0	V[Temp]	3.825	0.137	27.932	0
Daphnia <-> Daphnia	0	V[Daphnia]	1.218	0.07	17.472	0
Leptodora <-> Leptodora	0	V[Leptodora]	1.373	0.069	19.904	0
Cryptomonas <-> Cryptomonas	0	V[Cryptomonas]	0.953	0.041	23.043	0

Fig. S1 -- Graphical model for case study #2 involving Lotka-Volterra predator-prey dynamics (see Fig. 1 for details about graphical notation), log-abundance for prey lnx_t and predator lny_t, the Taylor-series approximation for abundance $\tilde{x}_t$ and $\tilde{y}_t$, a vector $ones$ representing a model intercept, and the four estimated interaction parameters $(\alpha, \beta, \gamma, \delta)$.

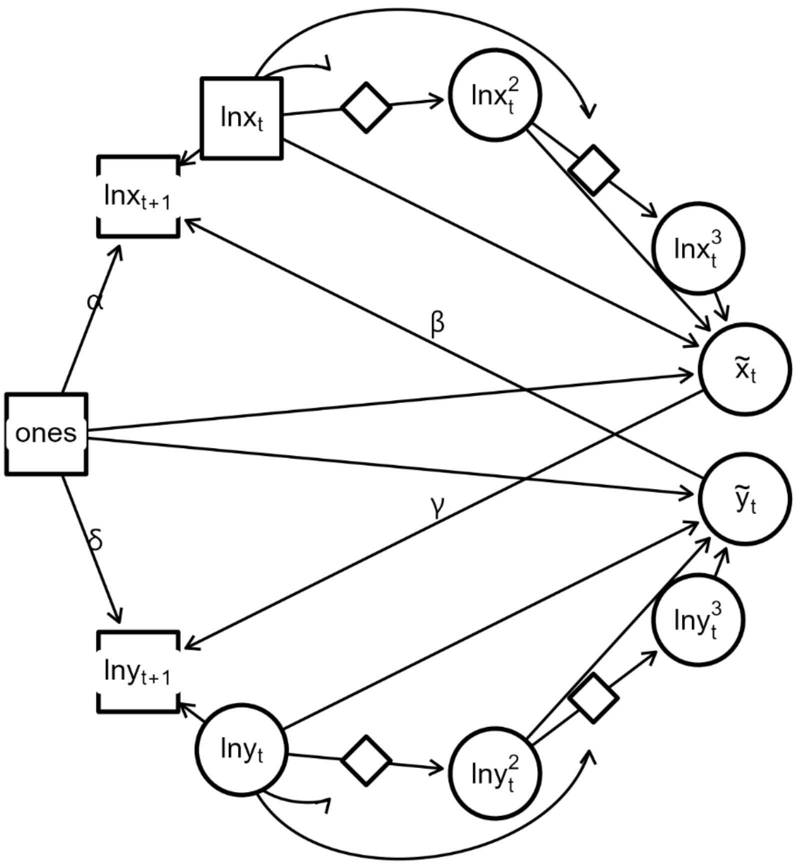

Fig. S2 – Observed values (circles), estimated values (lines), and 95% confidence intervals (shaded area) for case study #3.