nature ecology & evolution

Supplementary information

https://doi.org/10.1038/s41559-025-02864-8

Rigorous validation of ecological models against empirical time series

In the format provided by the authors and unedited

Supplementary Supplementary Information for

Rigorous validation of ecological models against empirical time series

Chuliang Song[1,2], Jonathan M. Levine[2]

[1] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA

[2] Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

Contents

Supplementary Note 1 Illustration of Covariance test in MacArthur-Rosenzweig model

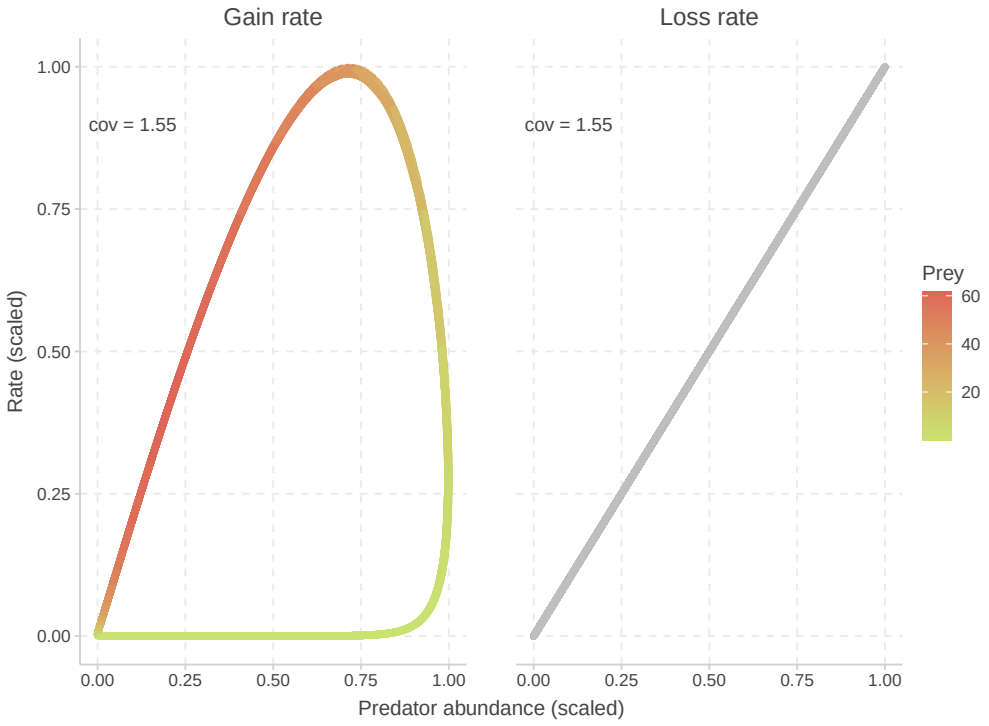
In the section "An Illustrated Worked-out Example" of the main text, we illustate the covariance criteria for Lotka-volterra dynamics. In this example, both rates are always higher when predator abundance is high, therefore creating a balance. Of course, ecological dynamics can exhibit more complex associations.

To establish the intuition why the covariance criteria holds for more complex scenarios, we consider the MacArthur-Rosenzweig predator-prey model with a Type II functional response:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = e\left(\frac{aN}{1+ahN}\right)P - dP, \tag{S1}$$

where N is the prey population size, P is the predator population size, r is the intrinsic growth rate of the prey, K is the carrying capacity of the environment for the prey, a is the attack rate of the predator, h is the handling time per prey item for the predator, e is the conversion efficiency of consumed prey into predator biomass, d is the natural death rate of the predator.

In this model, the gain rate of the predator depends non-linearly on both predator and prey abundances due to the Type II functional response. The figure below illustrates how gain and loss rates vary with predator abundance:



Supplementary Figure S1: Gain (left panel) and loss (right panel) rates for the predator P in the MacArthur-Rosenzweig model. The covariances between gain/loss rates and abundance are equal, satisfying the covariance criterion. The colors in left panel represent the prey abundance along the trajectories.

We see that the nonlinearity in the gain rates can be sepearted into two parts: when prey abundance N is low versus high. Although the gain rate does not always covary strongly with P (specifically when N is low), the periods when N is high contribute significantly to the overall covariance between gain rate and P. This is because the magnitude of the gain rate is much

larger during these periods. Thus, when considering the entire time series, the positive covariance during high N periods outweighs the weak covariance during low N periods. Thus, both gain and loss rates covary positively with predator abundance. As an illustrative example, we include an example using the MacArthur-Rosenzweig predator-prey model with a Type II functional response:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = e\left(\frac{aN}{1+ahN}\right)P - dP, \tag{S2}$$

where N is the prey population size, P is the predator population size, r is the intrinsic growth rate of the prey, K is the carrying capacity of the environment for the prey, a is the attack rate of the predator, h is the handling time per prey item for the predator, e is the conversion efficiency of consumed prey into predator biomass, d is the natural death rate of the predator.

Supplementary Note 2 Testing Discrimination Power of the Covariance Criteria

2.1 Compliation of 40 functional responses

To facilitate evaluation, we have reproduced Table 1 in Novak & Stouffer (2021), specifying each model's per-predator prey consumption rate as a function of prey abundance (N), predator abundance (P), and relevant parameters.

Table Supplementary Table 1: Functional-response models considered for describing the per-predator rate at which prey are eaten as a function of prey abundance N, predator abundance P, and the parameter(s).

Abbrevation	**Functional Response**	**Reference**
H1	aN	Lotka 1925; Volterra 1927
LR	$\frac{aN}{P}$	Pimm 1982; Arditi & Ginzburg 1989
BWL1	$\frac{a\sqrt{N}}{\sqrt{P}}$	Barbier *et al.* 2021
H2	$\frac{aN}{1+abN}$	Holling 1959
MM	$\frac{aN}{b+N}$	Michaelis *et al.* 1913
H3	$\frac{aN^2}{1+abN^2}$	Holling 1965; Real 1977
HT	$\frac{1}{b}\tanh(abN)$	Jassby & Platt 1976
GI	$\frac{1}{b}(1-\exp[-aN])$	Gause 1934; Ivlev 1955
GIA	$\frac{1}{b}(1-\exp[-abN])$	Aldebert *et al.* 2016a,b
GB	$\frac{1}{b}\left(1-\exp\left[-\frac{aN}{P}\right]\right)$	Gutierrez *et al.* 1984
A0	$\frac{aN}{1+ab\sqrt{N}}$	Abrams 1982
A1	$\frac{\sqrt{a}N}{1+abN}$	Abrams 1990
A3	$\frac{a\sqrt{N}}{1+ab\sqrt{N}}$	Abrams 1990
SH	$\frac{aN}{1+abN^2}$	Sokol & Howell 1981
AG	$\frac{aN/P}{1+abN/P}$	Sutherland 1983; Arditi & Ginzburg 1989
CDAO	$\frac{aN/\sqrt{P}}{1+abN/\sqrt{P}}$	Cosner *et al.* 1999
AGK	$\frac{a(N/P)^2}{1+ab(N/P)^2}$	Kratina *et al.* 2009
R	aN^u	Rosenzweig 1971
HV	$\frac{aN}{P^v}$	Hassell & Varley 1969
H3R	$\frac{aN^u}{1+abN^u}$	Real 1977
AS	$\left(\frac{aN}{1+abN}\right)^u$	Novak & Stouffer 2021
HLB	$\frac{aN^2}{1+cN+abN^2}$	Hassell *et al.* 1977
MH	$\frac{aN}{1+cN+abN^2}$	Andrews 1968
T	$\frac{aN}{1+abN+cN^3}$	Tostowaryk 1972
FHM	$\frac{aN\exp[dN]}{1+abN\exp[dN]}$	Fujii *et al.* 1986
A2	$\frac{aN}{1+abN+\sqrt{acN(1+abN)}}$	Abrams 1990
SSS	$\frac{2aN}{1+a(b+c)N+\sqrt{1+aN(2(b+c)+aN(b-c)^2)}}$	Jeschke *et al.* 2002
RGD	$\frac{2aN}{1+abN+\sqrt{(1+abN)^2+8ac(P-1)}}$	Ruxton *et al.* 1992; Cosner *et al.* 1999
BD	$\frac{aN}{1+abN+c(P-1)}$	Beddington 1975; DL 1975
CM	$\frac{aN}{1+abN+c(P-1)+abcN(P-1)}$	Crowley & Martin 1989
TTA	$\frac{aN}{1+abN+cP-(1-\exp[-cP])}$	Tyutyunov *et al.* 2008

Table Supplementary Table 1 – continued from previous page

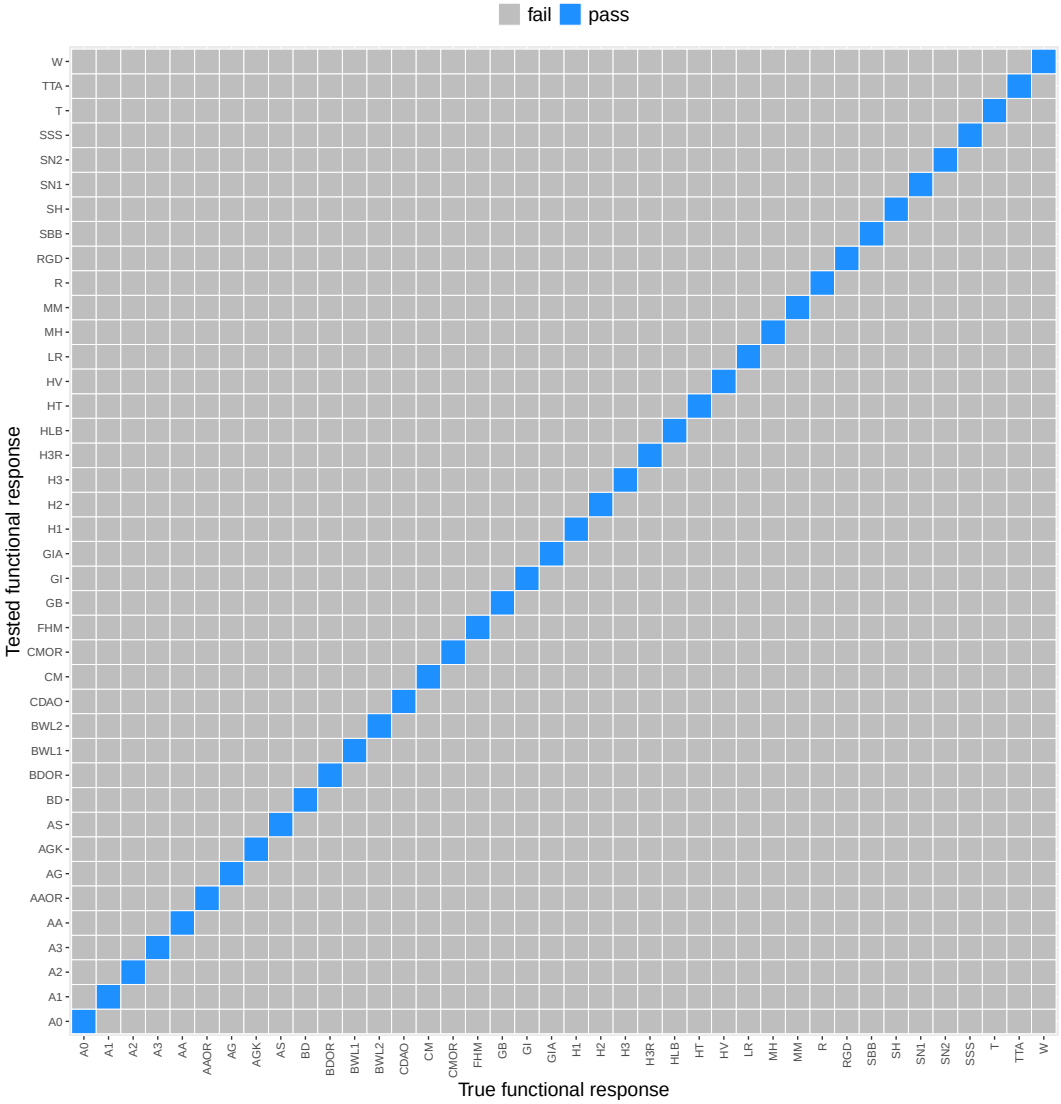
Abbrev.	Functional Response	Reference
BWL2	$aN^u P^{v-1}$	Barbier *et al.* 2021
AA	$\frac{aN/P^v}{1+abN/P^v}$	Arditi & Akçakaya 1990
SBB	$\frac{a(N/P^v)^2}{1+ab(N/P^v)^2}$	Schenk *et al.* 2005
W	$\frac{1}{b}\left(1-\exp\left[-\frac{aN}{P^v}\right]\right)$	Watt 1959
BDOR	$\frac{aN^u}{1+abN^u+c(P-1)}$	Okuyama & Ruyle 2011
CMOR	$\frac{aN^u}{1+abN^u+c(P-1)+abcN^u(P-1)}$	Okuyama & Ruyle 2011
AAOR	$\frac{aN^u/P^v}{1+abN^u/P^v}$	Okuyama & Ruyle 2011
SN1	$\frac{aN}{1+abN+c(P-1)+abc(1-d)N(P-1)}$	Stouffer & Novak 2021
SN2	$\frac{aN(1+c(1-d)(P-1))}{1+abN+c(P-1)+abc(1-d)N(P-1)}$	Novak & Stouffer 2021

2.2 Perfect Discriminatory Power without noise

We simulated the prey and predator dynamics for each of these functional responses. As shown in Figure S2, all models produce all models produce persistent cyclic behavior, making it challenging to differentiate between them.

Supplementary Figure S2: Each panel shows the dynamics under different functional response. Red denotes prey while blue denotes predator. All of them are in a stationary dynamics with persistent cyclic behavaor.

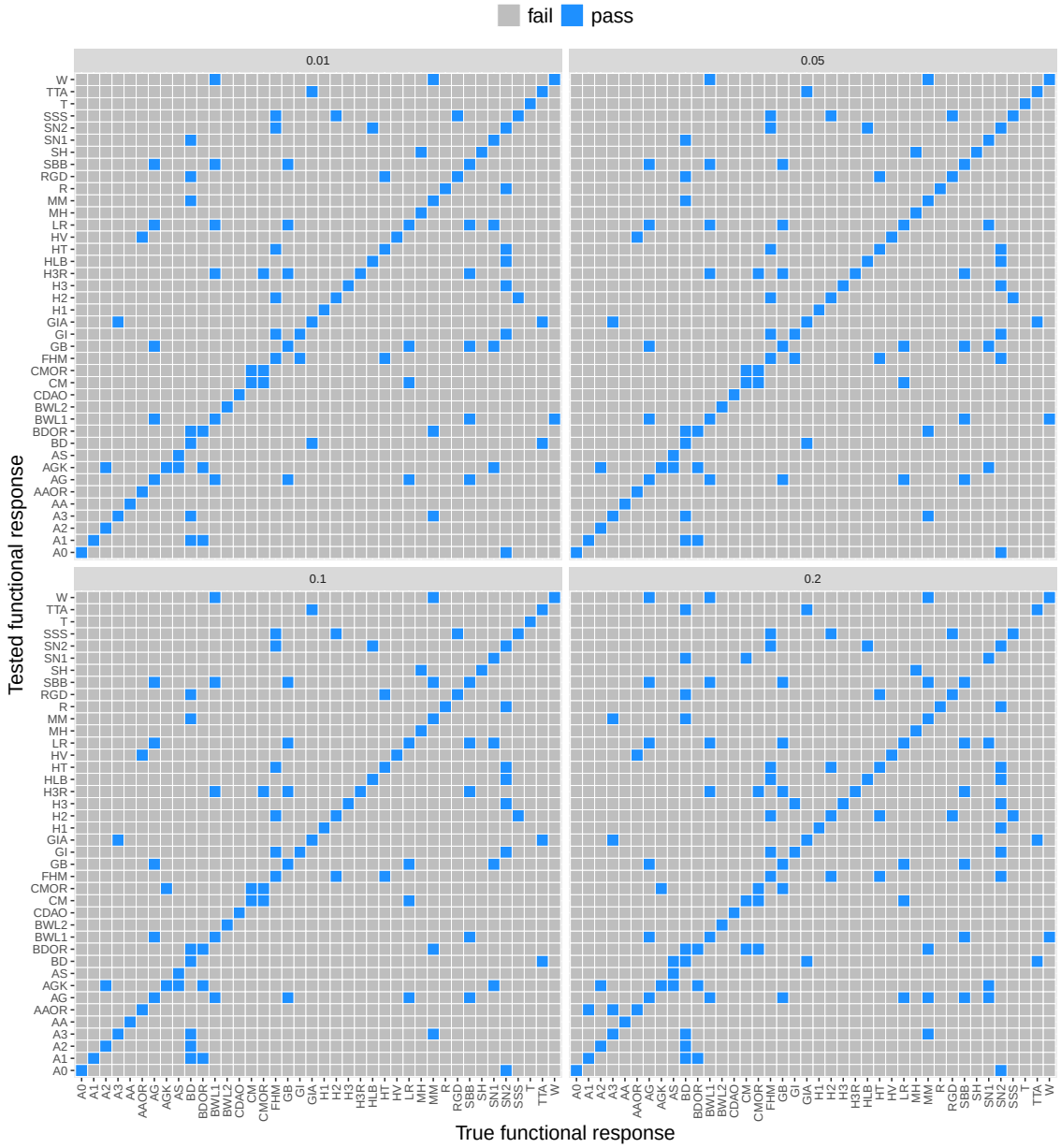
To assess the discriminative power of the covariance approach, we analyzed the covariance structures resulting from each model's dynamics. Our findings, presented in Figure S3, reveal that each functional response model exhibits a unique covariance pattern. This allows for clear differentiation among the models based on their covariance structures, even when their time series appear visually similar.



Supplementary Figure S3: Comparison of true and tested functional responses using the covariance test. Each cell represents a test of the time series in Figure S2 where the true response (x-axis) is compared to a tested response (y-axis). Blue indicates the tested response passed the covariance test, gray indicates failure. Response abbreviations are as defined in Table 1.

Supplementary Note 3 Type I vs Type II error of covariance criteria

An important caveat, though, is that the results above are obtained under idealized, noise-free conditions. To address potential real-world applications where measurement error or environmental noise might influence the results, we performed additional simulations incorporating varying levels of stochastic noise (ranging from 1% to 20%). As shown in Figure S4 below, the covariance approach successfully identified the correct model in most cases, even with the added complexity of noise:



Supplementary Figure S4: Same as Figure S3 expect with adding noise from 1% to 20%.

Supplementary Note 4 Applying covariance criteria to complex models

4.1 Logistic growth model

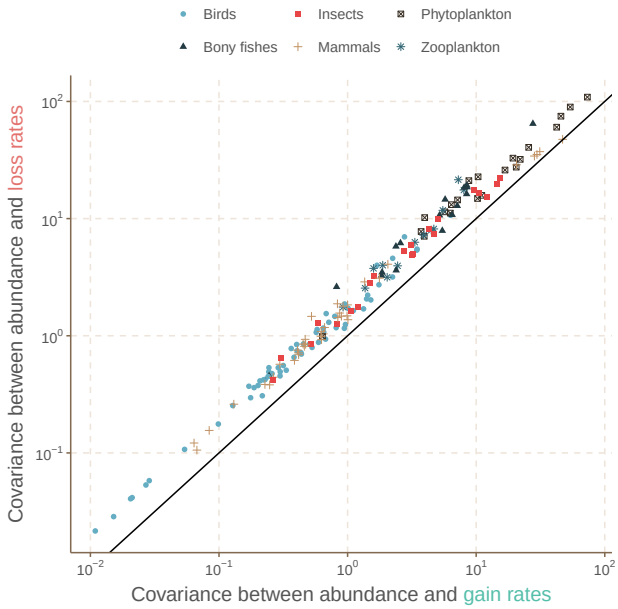
We first tested the classic logistic growth model:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN(1 - \frac{N}{K}) \tag{S3}$$

In this model, the covariance criteria are non-parametric because the parameters can be canceled out:

$$\frac{\mathtt{Cov}(\cancel{r}N, N)}{\langle \cancel{r}N \rangle \langle N \rangle} = \frac{\mathtt{Cov}(\cancel{r}/\cancel{K}N^2, N)}{\langle \cancel{r}/\cancel{K}N^2 \rangle \langle N \rangle} \tag{S4}$$

After canceling the parameters, we can test the model's predictions directly against the data. Despite the data being collected from a wide range of systems, almost none of the time series fall on the one-to-one line predicted by the model (Figure S5):



Supplementary Figure S5: Covariance analysis for the logistic growth model. Each point represents a different time series, and colors denote different taxonomic groups in the dataset. The black line represents the one-to-one line predicted by the model.

With statistical testing, we found that only 10 out of 172 time series pass the covariance test. Since the z-score threshold ($= 1.98$) corresponds to a p-value less than 0.05, the proportion $10/172 \approx 0.058$ is roughly what would be expected by chance alone. This suggests that the logistic growth model does not adequately describe the dynamics of most populations in this dataset.

4.2 Theta logistic model

Next, we analyzed the theta logistic model (Gilpin & Ayala, 1973; Thomas *et al.*, 1980):

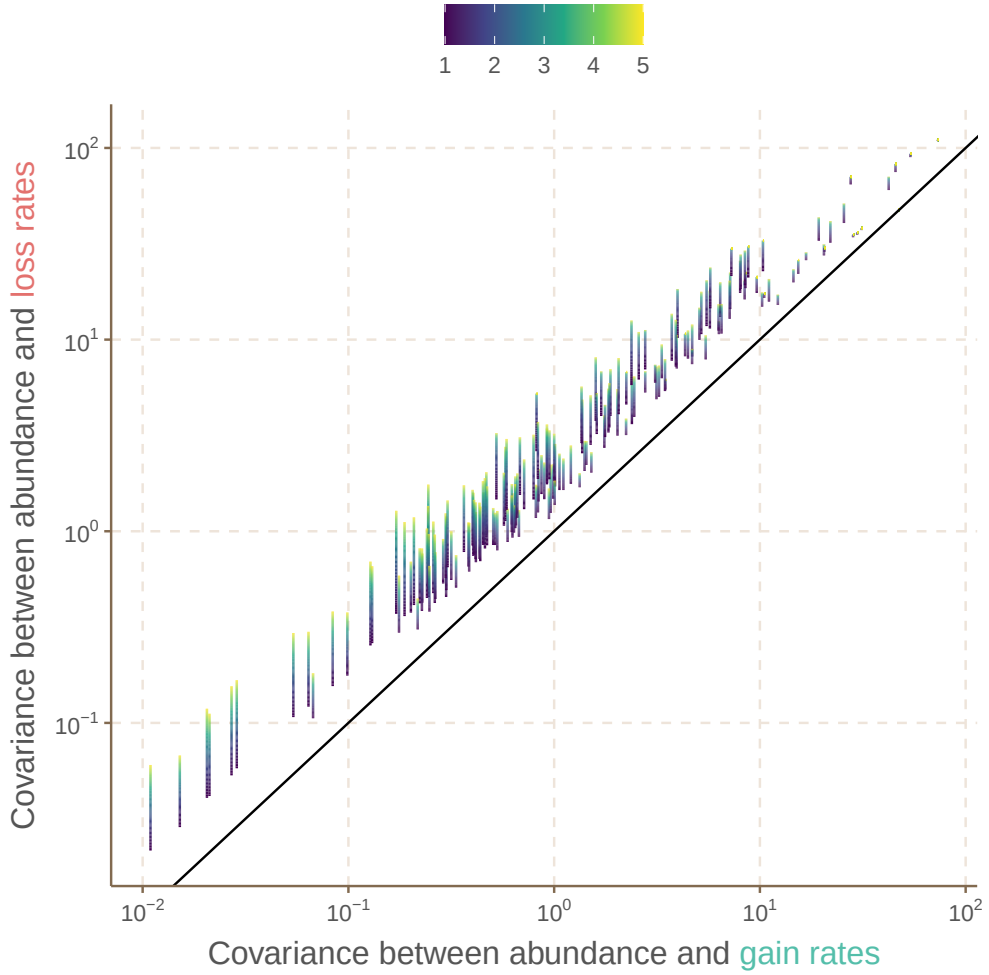
$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN\left(1 - (\frac{N}{K})^{\theta}\right) \tag{S5}$$

where $\theta > 1$. When $\theta = 1$, it reduces to the logistic growth model.

The covariance criterion is no longer non-parametric due to the unknown θ:

$$\frac{\texttt{Cov}(\cancel{r}N, N)}{\langle \cancel{r}N \rangle \langle N \rangle} = \frac{\texttt{Cov}(\cancel{1/K^\theta} N^{1+\theta}, N)}{\langle \cancel{1/K^\theta} N^{1+\theta} \rangle \langle N \rangle} \tag{S6}$$

We varied θ from 1 to 5, as it is required to be larger than 1 to ve biologically meaningful. Figure S6 shows the covariance test results:



Supplementary Figure S6: Covariance test results for the theta growth model across varying values of θ.

As θ increases, the covariance with the loss rate deviates further from that with the gain rate. Thus, despite the unknown parameter, the covariance analysis allows us to confidently reject the theta growth model as an explanation for the observed data.

4.3 Sublinear growth model

We have tested this on recently proposed sublinear growth model (Hatton *et al.*, 2024; Mazzarisi & Smerlak, 2024):

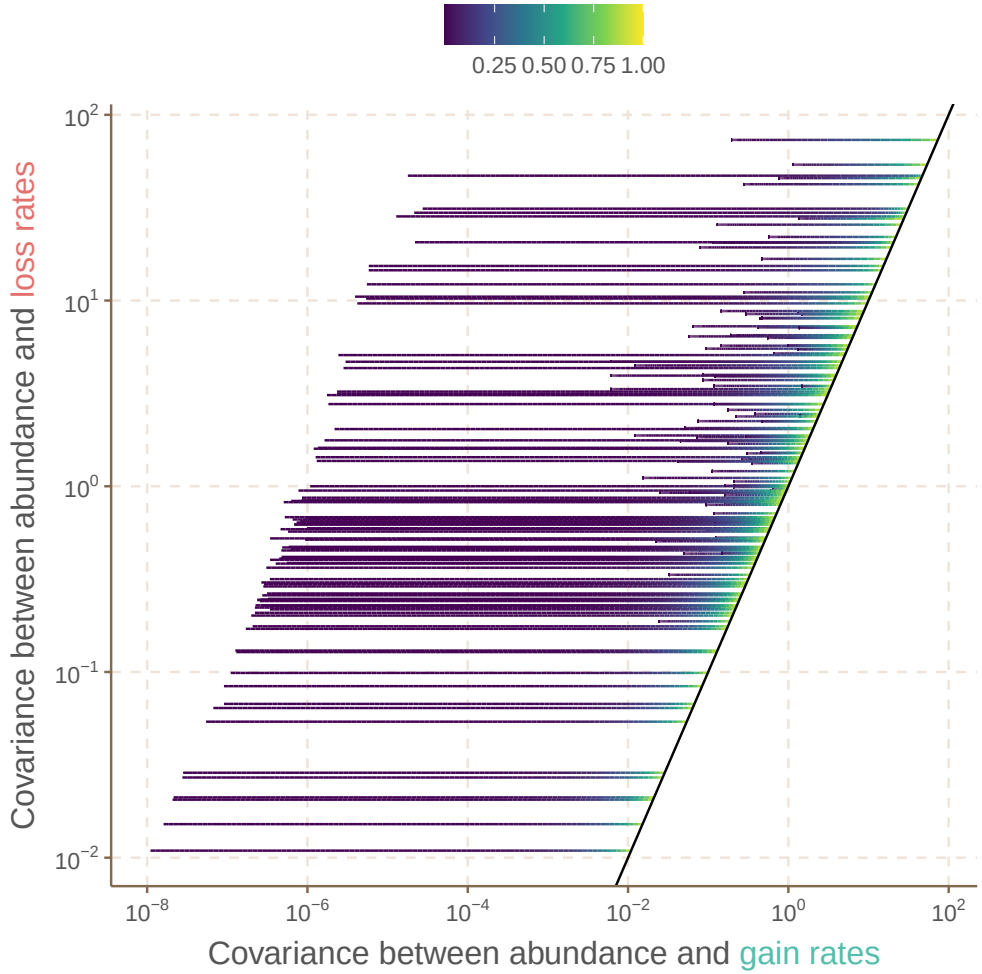
$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN^k - bN \tag{S7}$$

where $0 < k < 1$. The parameter k accounts for the sublinear growth. Note that this model cannot be reduced to the logistic growth model (there is no possibility of an N^2 term).

The covariance criterion is also not non-parametric due to the unknown k:

$$\frac{\texttt{Cov}(\not{r}N^k, N)}{\langle \not{r}N^k \rangle \langle N \rangle} = \frac{\texttt{Cov}(\not{b}N, N)}{\langle \not{b}N \rangle \langle N \rangle} \tag{S8}$$

We systematically varied k from 0 to 1:



Supplementary Figure S7: Covariance test results for the sublinear growth model across varying values of k.

At first glance, it may seem that the model predictions approach the one-to-one line and pass the covariance test when k is very close to 1. However, this is an artifact. The system only appears to pass the covariance test when it loses its sublinear character (i.e., when $k \approx 1$). Additionally, as we discussed in response to Reviewer 1's insightful comment below, the covariance criteria do not apply when the gain and loss rates have the same functional form (here, $k = 1$) because there will be no fltuating dynamics (please refer to our detailed reply below). Thus, we can still confidently reject model with the covariance criteria.

The results above are not surprising. Most populations do not exist in isolation from other species; it is simply that we only have data on one speciesc in this global dataset. However, using the covariance criteria, we have made this verbal argument rigorous. As a practical application, this results warns the danger of working with this dataset without accounting for the effects of unobserved species (Pennekamp *et al.*, 2019; Rogers *et al.*, 2022).

Supplementary Note 5 Covariance criteria and the nature of fluctuations

A core assumption of the covariance criteria is that observed population fluctuations are primarily driven by the deterministic interplay of gain and loss processes, rather than being completely overwhelmed by stochastic noise. To illustrate the boundaries of this assumption, we systematically test the method's performance in the presence of three canonical types of stochasticity.

We use the MacArthur-Rosenzweig predator-prey model as our testbed. This model generates a stable limit cycle due to prey density-dependence (logistic growth) and a saturating predator functional response (Type II). This is in direct contrast with the Lotka-Volterra predation model, where the rapid extinctions can easily occur in the presence of stochastic noise. This allows for a more meaningful analysis of how sustained stochasticity affects the covariance criteria.

The baseline deterministic MacArthur-Rosenzweig model is:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN\left(1 - \frac{N}{K}\right) - \frac{aNP}{1 + ahN} \tag{S9}$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = e\frac{aNP}{1 + ahN} - mP \tag{S10}$$

where N and P are prey and predator abundances, r is the prey's intrinsic growth rate, K is its carrying capacity, a is the predator's attack rate, h is its handling time, e is the conversion efficiency, and m is the predator mortality rate.

We introduce three different forms of stochastic noise to this system and observe how the proportion of failed validation tests changes as the noise intensity (σ) increases. The results are summarized in Figure S8.

5.1 Additive Noise

Additive noise represents fluctuations independent of population size, such as from constant environmental perturbations or measurement error. The stochastic term added to the prey and predator dynamics is σdW_t.

As theoretically expected, the covariance criteria are exceptionally robust to additive noise. The noise term is uncorrelated with population size, so it does not interfere with the covariance structure of the underlying deterministic dynamics. For both prey and predator, the proportion of failed tests remains near zero across all tested noise intensities (Figure S8, left panels).

5.2 Multiplicative (Environmental) Noise

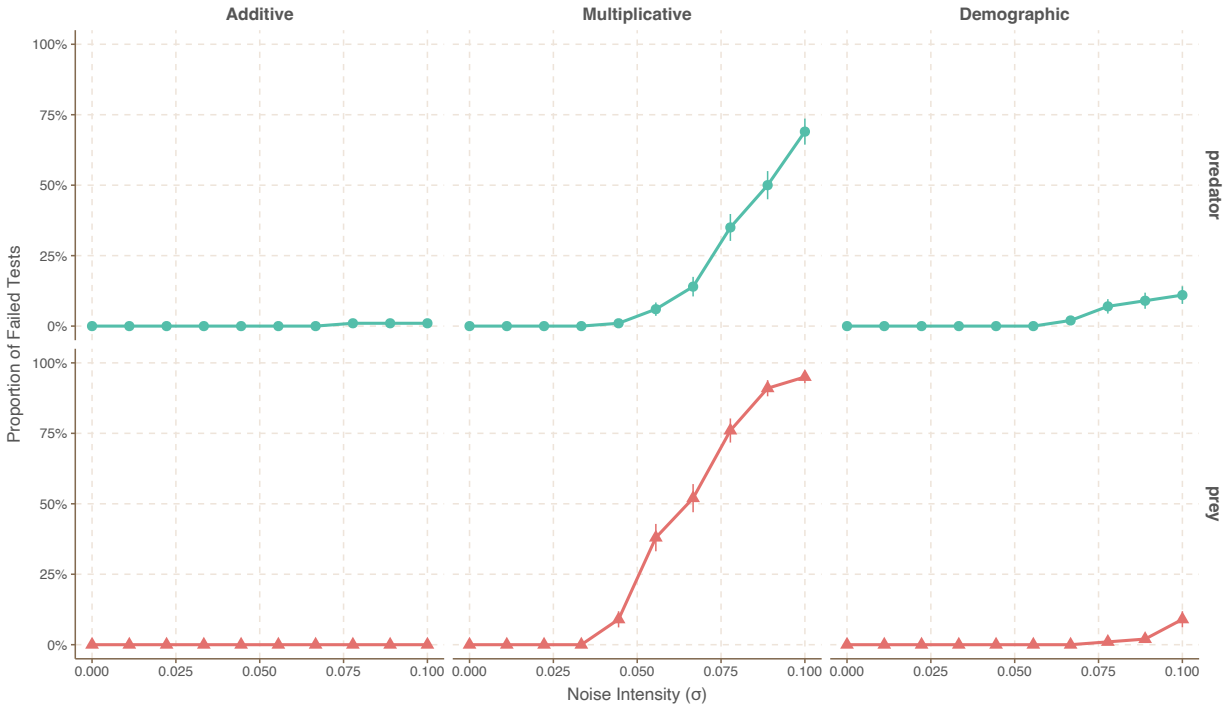
Multiplicative noise represents environmental fluctuations that affect per-capita demographic rates (e.g., a good year increasing r). This common form of density-dependent noise is modeled by adding a stochastic term of the form $\sigma N dW_t$ to the prey dynamics and $\sigma P dW_t$ to the predator dynamics.

Here, the noise term is correlated with population size and thus contributes to the overall covariance. When this noise becomes strong enough to rival the deterministic signal, the test begins to fail. Our simulations show this clearly: the proportion of failed tests for both species increases sharply as the noise intensity σ rises, approaching 100% failure at higher noise levels (Figure S8, middle panels).

5.3 Demographic Noise

Demographic noise arises from the inherent randomness of individual birth and death events. The standard deviation of this noise scales with the square root of the population size, modeled with a stochastic term of the form $\sigma\sqrt{N}dW_t$.

Like multiplicative noise, demographic noise is density-dependent and can disrupt the covariance criteria. However, because its magnitude scales more slowly with population size ($\sqrt{N}$ vs. N), its impact is less severe at equivalent σ values. The proportion of failed tests increases with noise intensity but at a much slower rate than for multiplicative noise, particularly for the predator (Figure S8, right panels). This demonstrates that the stabilizing features of the MacArthur-Rosenzweig model make it resilient to moderate levels of demographic stochasticity.



Supplementary Figure S8: **Performance of the covariance criteria for the MacArthur-Rosenzweig model under different noise structures.** Each panel shows the proportion of simulations that failed the covariance test (z-score > 1.96) as a function of increasing noise intensity (σ). Rows correspond to the predator and prey, while columns represent the type of stochastic noise. Error bars denote the standard error of the proportion. The method is robust to additive noise but sensitive to high levels of density-dependent (multiplicative and demographic) noise, with the effect being most pronounced for multiplicative noise.

These simulation results confirm that the validity of the covariance criteria depends crucially on the signal-to-noise ratio. The method is robust to additive noise but will be compromised when density-dependent noise (multiplicative or demographic) is strong enough to overwhelm the deterministic dynamics. This is not a failure of the method, but rather a fundamental limit of any approach attempting to infer deterministic rules from highly stochastic data.

Supplementary Note 6 Scenarios Where Covariance Criteria Do Not Apply

The covariance criteria, while powerful, do have limitations in their applicability. Here, we outline scenarios where the method might not be suitable or requires careful consideration.

6.1 System without fluctuation

Our approach is specifically designed for systems exhibiting fluctuating dynamics—situations where gain and loss rates vary over time, leading to alternating periods dominated by gains and losses. In models where gain and loss rates are governed by identical functions, the net change would lack the necessary fluctuations for our covariance criteria to be meaningfully applied.

To illustrate this, consider a population model described by identical functional forms of gain and loss rates:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN - dN \tag{S11}$$

where where r is the per capita birth rate and d is the per capita death rate:

- If $r = d$: The net growth rate is zero, and the population remains at equilibrium indefinitely, with no fluctuations occurring. In this scenario, our method cannot extract meaningful insights because fluctuations are essential for covariance analysis. This is not a limitation unique to our approach; any method would struggle to reveal underlying dynamics in a system that remains perpetually at equilibrium.
- If $r \neq d$, The system would either grow exponentially or decline to extinction, again without the necessary fluctuations in gain and loss rates over time.

Therefore, our approach would not erroneously validate models where gain and loss rates follow identical functions, as the requisite dynamic variability is absent.

6.2 Ecological systems with strong trends

The covariance criteria are rigorously applicable to stationary systems, where the statistical properties of the system remain constant over time, or to cyclostationary systems, where statistical patterns repeat predictably (e.g., with seasons). We have tested this on the empirical datasets used in this paper, and they are mostly stationary.

In this context, it is worth noting that many ecological systems can exhibit non-stationary dynamics. A prime example is populations experiencing consistent growth or decline, which violates the stationarity assumption. In such cases, applying the covariance criteria directly to the entire time series might lead to misleading results.

As a simple illustrative example, consider the logistic growth model:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = rN(1 - N/K) \tag{S12}$$

We simulate this model with parameters $r = 0.5$ and $K = 1000$. Applying the covariance criteria to this simulated data results in a z-score greater than 7. This indicates that the covariance with gain and the covariance with loss are significantly different, which is expected given that, at least initially, logically growing systems have a higher gain than loss rate.

6.3 Unobserved Direct Interactions or Frequent Directional Perturbations

To effectively utilize the covariance criteria, it is crucial to have knowledge of all the key variables influencing the gain and loss terms, particularly any direct interactions with other species or significant external perturbations. The covariance criteria assume that external perturbations, such as environmental fluctuations, are statistically independent of the population abundance. If there are significant unobserved direct interactions or frequent and strong external perturbations that correlate with population abundance, the gain and loss rates are influenced by unknown factors, and the method might encounter difficulties in differentiating between various models or identifying model inadequacies.

6.4 Fluctuation driven by noise instead of deterministic dynamics

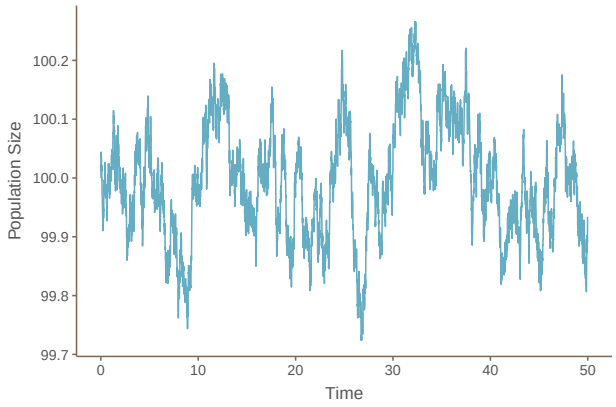
As we detailed in Appendix 5, our covariance criterion assumes that fluctuations in population size are primarily driven by the inequality between gain and loss rates. As a concrete example, we show the applicability of the covariance criterion varies based on the type of stochasticity:

- **Fluctuations Driven by White Noise**. When fluctuations are driven purely by white noise (represented by additive stochasticity), the correlation between the gain rate and the population size is effectively zero, as is the correlation between the loss rate and the population size. In such cases, both sides of the covariance criterion are zero, providing no useful information.

  As a concrete example, consider the stochastic logistic growth model with additive white noise:

  $$dN = rN(1 - N/K)dt + \sigma dW \tag{S13}$$

  We ran a simulation of this model:



Supplementary Figure S9: Time series of population size N with fluctuations driven by white noise.

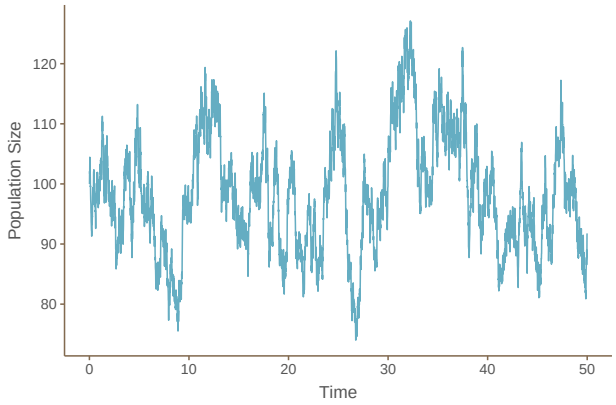
  In this case, the normalized covariance between gain and abundance is 10^{-8}, and that between loss and abundance is 10^{-7}—both effectively zero. Therefore, the covariance criterion does not provide meaningful information in this scenario, as predicted.

- **Fluctuations Driven by Density-Dependent Noise**. When fluctuations are driven by density-dependent noise, such as demographic and multiplicative stochasticity, this can introduce correlations between the gain/loss rates and the population size. However, these correlations arise from the noise itself rather than from the deterministic dynamics, potentially leading to misleading conclusions when applying the covariance criterion.

As a concrete example, consider the stochastic logistic growth model with environmental noise:

$$dN = rN(1 - N/K)dt + \sigma N dW \tag{S14}$$

We ran a simulation of this model using the same random seed as before:



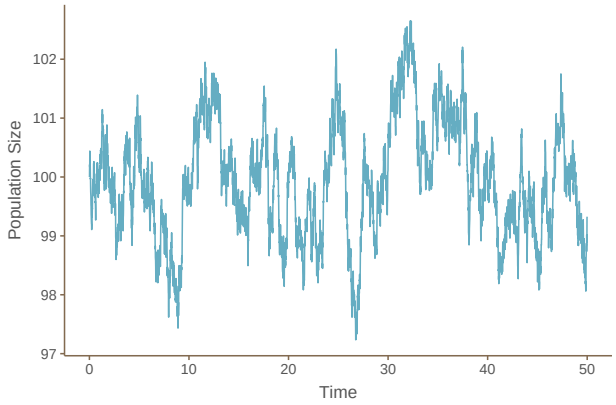
Supplementary Figure S10: Fluctuations driven by density-dependent environmental noise.

Although the time series appears visually similar to the previous case, the covariance structure is different. The normalized covariance between gain and abundance is 0.01, and that between loss and abundance is 0.02, resulting in a high z-score of 45. These non-zero covariances are induced by the noise term rather than the deterministic gain and loss processes. Therefore, applying the covariance criterion without accounting for the nature of the noise may lead to incorrect conclusions.

As another example, consider the stochastic logistic growth model with demographic noise:

$$dN = rN(1 - N/K)dt + \sigma\sqrt{N}dW \tag{S15}$$

We ran a simulation of this model using the same random seed as before:



Supplementary Figure S11: Fluctuations driven by density-dependent demographic noise.

Similarly, although the time series appears visually similar to the previous case, the covariance structure is different (z-score = 19).

- **Fluctuations Driven by Deterministic Components (e.g., Seasonal Variations)**. When fluctuations are influenced by deterministic components, such as seasonal changes in carrying capacity, the covariance criterion remains applicable. In these cases, the

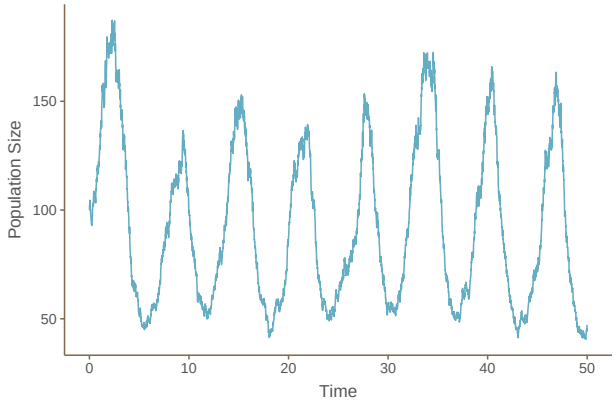
correlation between gain/loss rates and population size is meaningful and reflects the underlying deterministic dynamics.

Consider the stochastic logistic growth model with a periodically varying carrying capacity:

$$dN = rN(1 - N/K)dt + \sigma N dW \tag{S16}$$

where $K \propto \exp(\sin(t))$, introducing seasonal fluctuations in the carrying capacity. This choice is simply for illustrative purpose. Although the noise is density-dependent (due to the N term in $\sigma N dW$), it is relatively negligible compared to the deterministic seasonal effects.

We ran a simulation of this model:



Supplementary Figure S12: Fluctuations driven by deterministic seasonal variations.

In this example, the system passes the covariance test with a z-score of 0.33, demonstrating that the covariance approach can be effectively used in single-species models subjected to stochasticity around equilibrium when deterministic factors influence the fluctuations. The key is that deterministic components (e.g., seasonal changes in carrying capacity) maintain a meaningful correlation structure between gain and loss rates and the population size.

In light of these analyses, let us revisit the widespread failure of the logistic model demonstrated in Appendix 4.1. It can be interpreted in two ways:

- Interpretation A: The Model is Structurally Wrong. The most direct interpretation is that the fundamental assumptions of the logistic model (i.e., linear gain and purely quadratic loss) do not correctly describe the deterministic forces governing these real populations. The systems are likely regulated by more complex interactions, different forms of density dependence, or external drivers not included in the model.
- Interpretation B: The Dynamics are Noise-Dominated. Alternatively, as outlined in F.4, it is possible that for some of these populations, the logistic model is a reasonable approximation of the deterministic dynamics, but the observed fluctuations are so heavily dominated by environmental or demographic stochasticity that the deterministic signal is lost. In this scenario, while the model's deterministic part might be "correct," it is dynamically irrelevant for explaining the observed variance, and thus the model as a whole is still not a useful tool for understanding that system's fluctuations.

In either case, the conclusion is the same: the standard logistic model, when applied to these empirical time series, is not a valid or useful representation of the observed dynamics. The failure of the test in D.1 is therefore a robust result, and F.4 provides a deeper understanding of the potential reasons for such a failure.

6.5 Covariance criteria does not apply to non-linear transformed dynamical equation

Log-transformations are commonly used in ecological analyses due to their ability to linearize multiplicative processes and stabilize variances. However, in the context of our covariance criteria framework, log-transformations introduce fundamental challenges.

Consider the continuous-time Lotka-Volterra prey dynamics:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = rx - axy \tag{S17}$$

For these dynamics, our covariance criterion is given by:

$$\frac{\mathtt{Cov}(x, rx)}{\langle x\rangle\langle rx\rangle} = \frac{\mathtt{Cov}(x, axy)}{\langle x\rangle\langle axy\rangle} \tag{S18}$$

This criterion is derived under the assumption of **additive** stochastic events in population sizes Hilfinger *et al.* (2016). Specifically, we represent the continuous dynamics using discrete stochastic events:

$$x \xrightarrow{\frac{rx}{rx+axy}} x + \delta \text{ and } \; x \xrightarrow{\frac{axy}{rx+axy}} x - \delta \tag{S19}$$

Here, δ represents a small increment, while $\frac{rx}{rx+axy}$ and $\frac{axy}{rx+axy}$ represent the probabilities per unit time of gain and loss rates, respectively. For the discrete dynamics, the covariance criteria derived in Hilfinger *et al.* 2016 reads at:

$$\frac{\mathtt{Cov}(x, rx)}{\langle x\rangle\langle rx\rangle} = \frac{\delta}{\langle x\rangle} + \frac{\mathtt{Cov}(x, axy)}{\langle x\rangle} \tag{S20}$$

As δ approaches zero (considering infinitesimally small changes, especially since x typically represent density), the term $\frac{\delta}{\langle x\rangle}$ becomes negligible, and we recover our covariance criterion (Eqn. S18) in the continuous limit.

When we apply a log-transformation to the variables ($\tilde{x} = \mathsf{log}(x), \tilde{y} = \mathsf{log}(y)$), the prey dynamics become:

$$\frac{\mathrm{d}\tilde{x}}{\mathrm{d}t} = r - a\exp(\tilde{y}) \tag{S21}$$

While this transformation linearizes the multiplicative process and simplifies the equation, it fundamentally alters the nature of the discrete stochastic events. The corresponding discrete events for the log-transformed variable is:

$$\tilde{x} \xrightarrow{\frac{r}{r+a\exp(\tilde{y})}} \tilde{x} + \delta_0 \text{ and } \; \tilde{x} \xrightarrow{\frac{a\exp(\tilde{y})}{r+a\exp(\tilde{y})}} \tilde{x} - \delta_0 \tag{S22}$$

Here, δ_0 is some fixed increment in $\tilde{x}$. In terms of the original variable x, these events correspond to **multiplicative changes**:

$$x \xrightarrow{\frac{rx}{rx+axy}} e^{\delta_0}x \text{ and } \; \tilde{x} \xrightarrow{\frac{rx}{rx+axy}} e^{-\delta_0}\tilde{x} \tag{S23}$$

Although the transition probabilities remain the same in both models, the nature of the events differs fundamentally. This distinction is crucial because the covariance criterion established in Hilfinger *et al.* (2016) **only applies to additive** discrete events, not multiplicative or other nonlinear ones. Therefore, we must carefully consider how our continuous model maps onto discrete events. Ecological studies indicate that changes in species abundance x are typically additive, such as individual births and deaths. In contrast, while treating $\mathsf{log}(x)$ as an additive

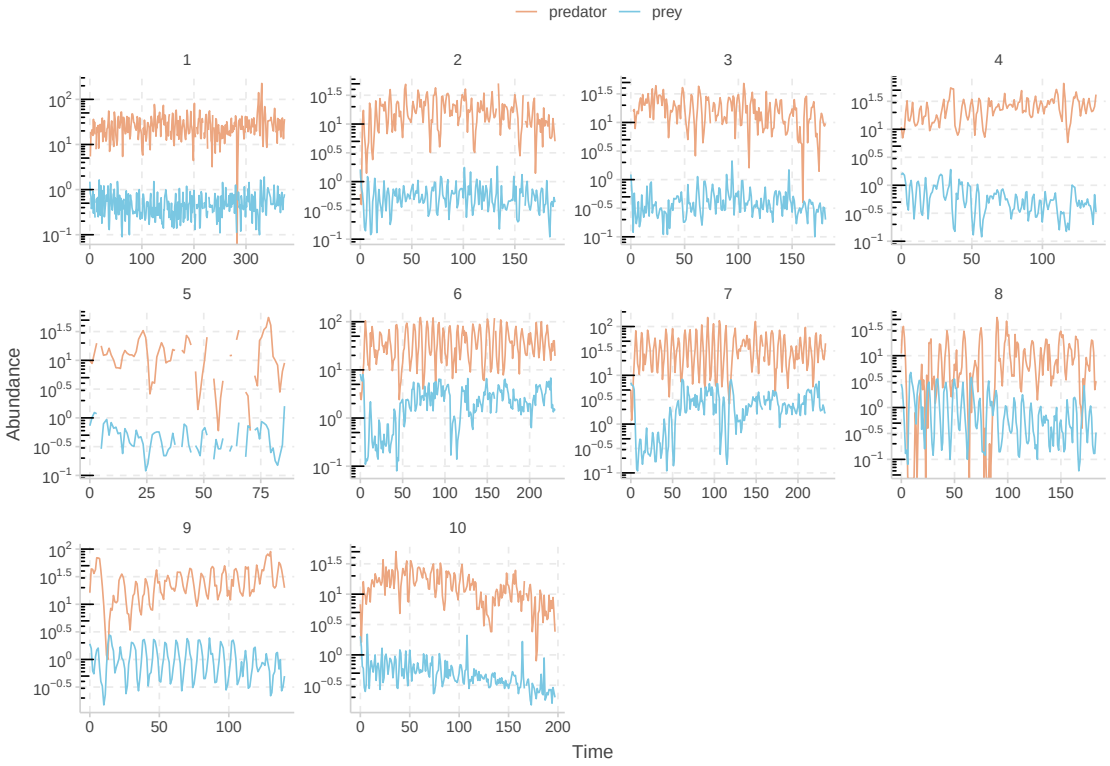
variable is mathematically feasible, it is not ecologically meaningful because it implies that x changes multiplicatively, which does not align with the ecological processes.

It might be unclear why these differences matter if both processes lead to the same continuous dynamics. An analogy may help clarify this point. Consider two functions, $f(x) = 0$ and $g(x) = x$, both of which equal zero when $x = 0$. However, their derivatives at $x = 0$ differ significantly ($f'(0) = 0$ for the constant function and $g'(0) = 1$ for the linear function), leading to different behaviors near $x = 0$. Similarly, we are examining the limiting case of the covariance criterion, not just the limiting behavior of the dynamics. Specifically, while the continuous dynamics remain the same after a log-transformation, the underlying discrete events—and thus the infinitesimal fluctuations relevant to our covariance criterion—are fundamentally different.

In conclusion, while the covariance criteria offer a powerful and versatile tool for model validation, it's essential to be mindful of its limitations and apply it judiciously, as is the case for any computational method.

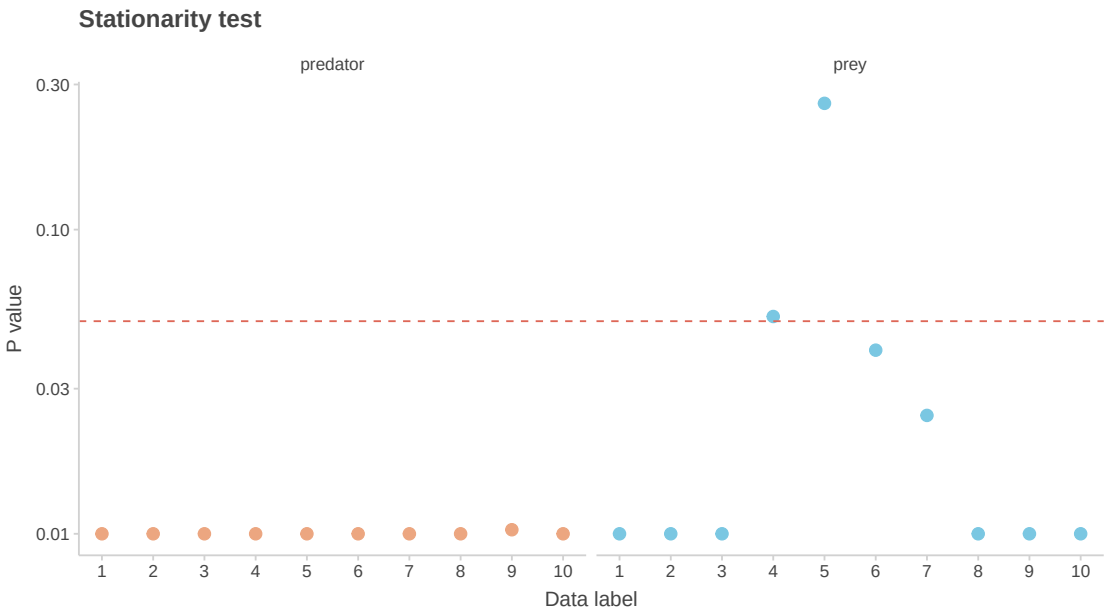
Supplementary Note 7 Reverse engineering the nature of predation

7.1 Visualizing raw data



Supplementary Figure S13: Time series of prey and predator abundances in 10 replicates from Blasius *et al.* (2020).

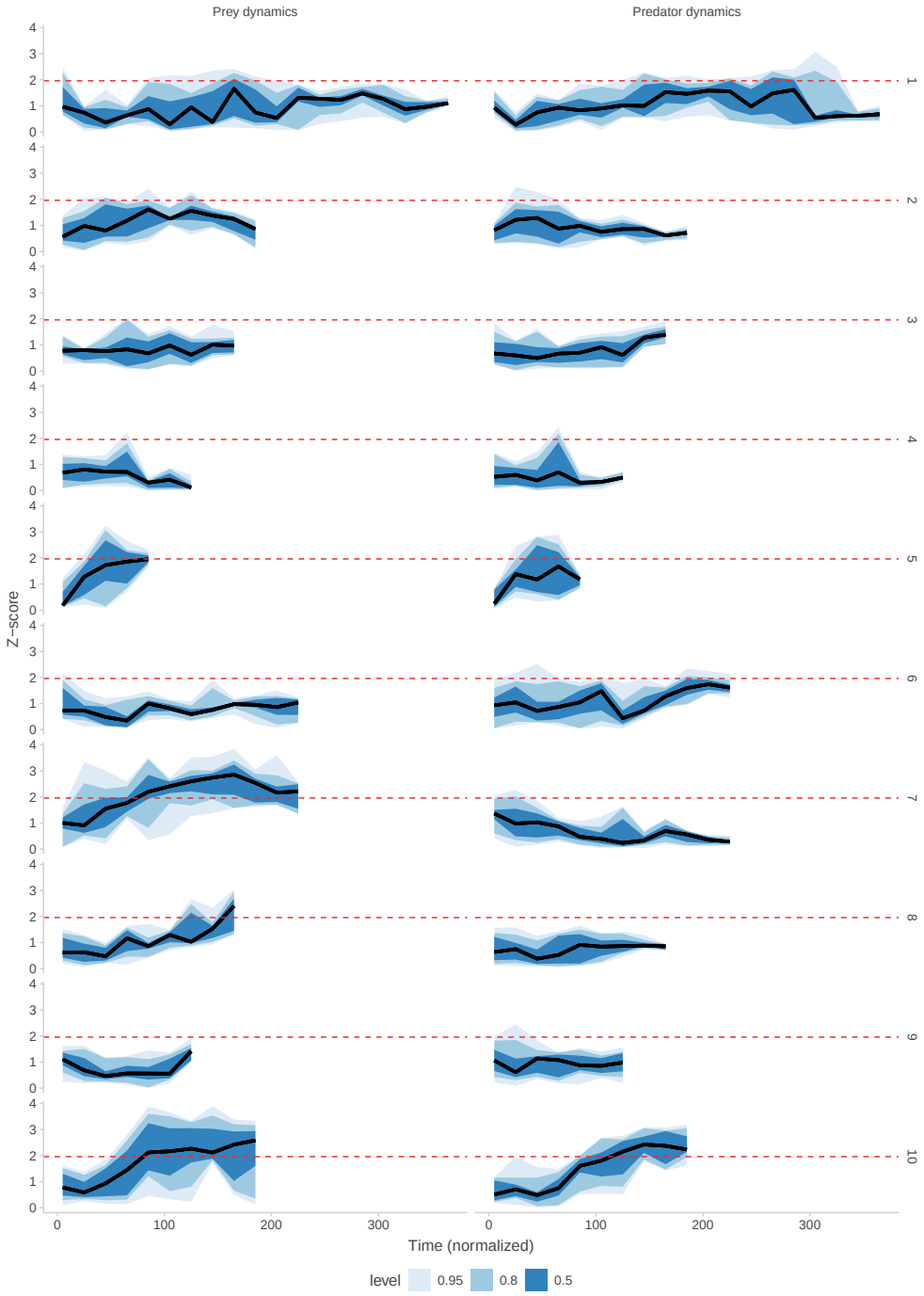
7.2 Stationarity test



Supplementary Figure S14: Stationarity test using augmented Dickey-Fuller test. P value less than 0.05 indicates evidence for stationarity.

7.3 Moving-window analysis

We conducted a moving window analysis on the ten replicates from Blasius *et al.* (2020). We varied the window length from 5 up to the maximum length of each time series. Within each window, we computed the z-score of the covariance test.



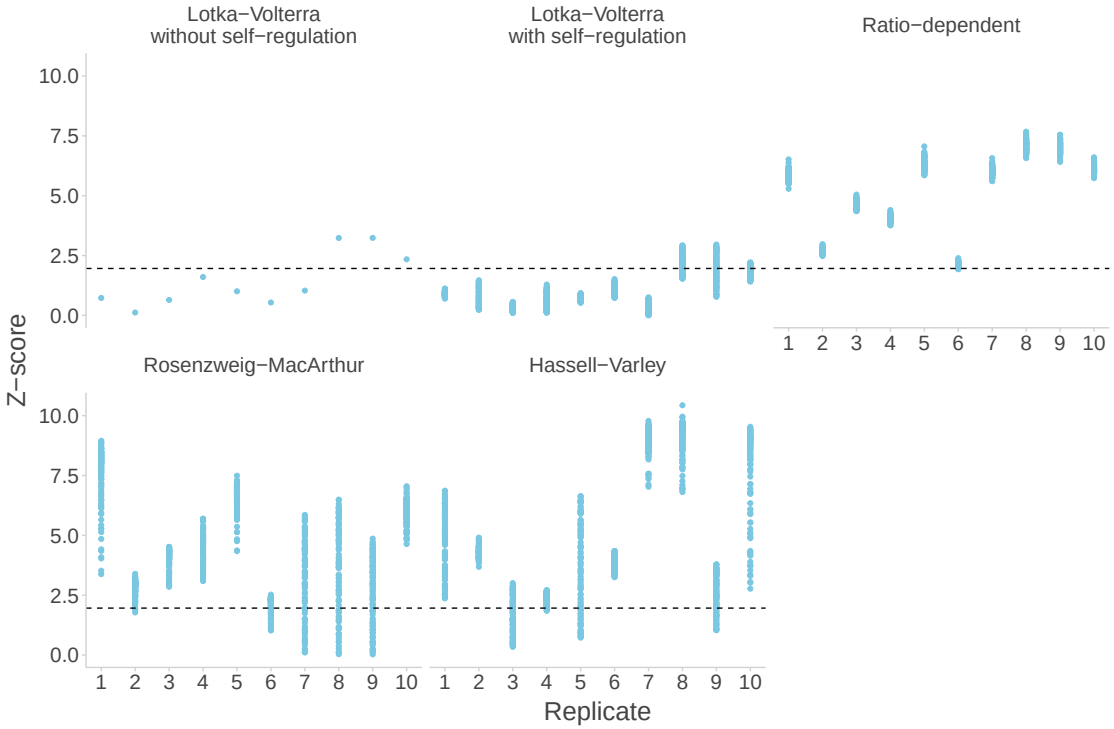
Supplementary Figure S15: **Moving window analysis with Lotka-Volterra dynamics**. The analysis was performed on the ten replicates from Blasius *et al.* (2020). The left panel shows results for the prey dynamics, while the right panel shows results for the predator dynamics. The x-axis represents the window length, and the y-axis displays the z-score of the covariance test. The black line indicates the average z-score across windows, and the shaded areas (ribbons) represent the 50%, 80%, and 95% confidence intervals.

As shown in Figure S15, the z-scores remain relatively constant across different window lengths for most replicates. This consistency suggests that our z-score approach is robust to variations in sample size within the tested range. Therefore, we believe that our results are unlikely to be an artifact of short time series lengths.

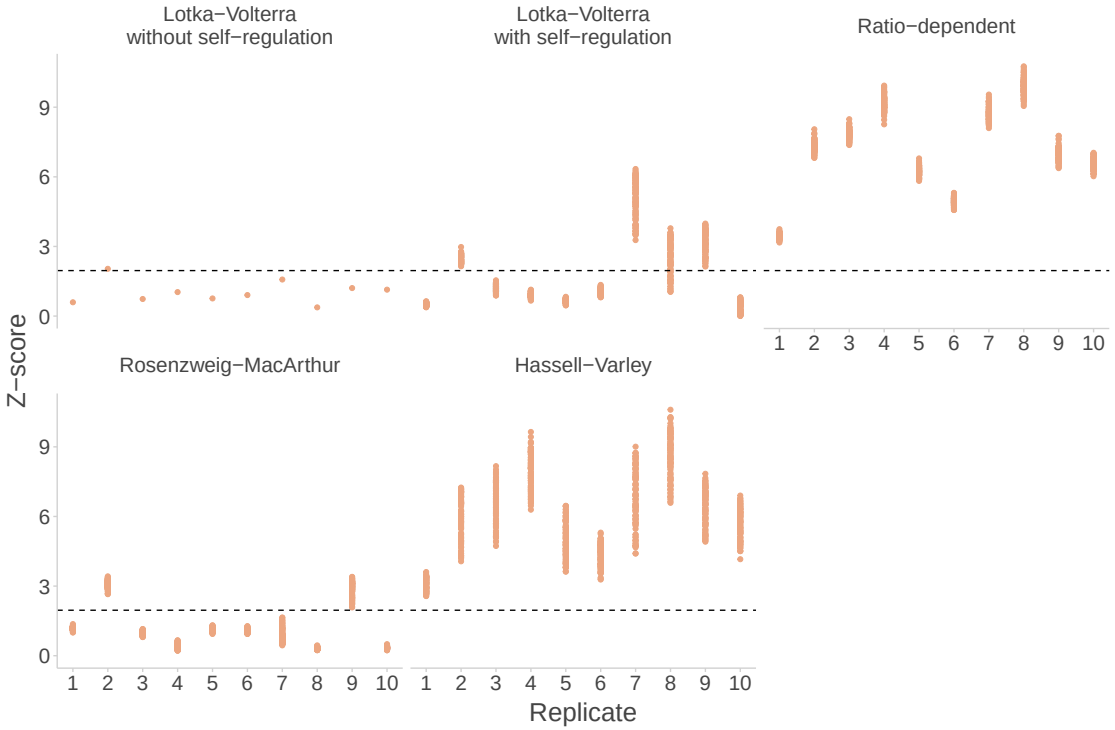
7.4 Testing alternative models

To conclusively determine whether ratio dependence is the source of the model's invalidity, it is essential to include both a saturating prey-dependent model and a non-saturating ratio-dependent model in our analysis.

To address this, we have expanded our analysis to include the Rosenzweig-MacArthur model (a saturating prey-dependent model) and the Hassell-Varley model (a non-saturating ratio-dependent model) using the datasets from Blasius *et al.* (2020). Following our solution to uncancellable parameters (Appendix 4), we systematically varied the parameters of these models spanning one order of magnitude to thoroughly test their validity. Figures S16 and S17 show our findings.



Supplementary Figure S16: Covariance test for prey dynamics across ten experimental replicates.



Supplementary Figure S17: **Covariance test for predator dynamics**. Same as Figure S16 expect we test the predator dynamics.

To intrepret the results:

- **Rosenzweig-MacArthur model**. The analysis shows that Rosenzweig-MacArthur model captures the predator dynamics well (Figure S17), but not the prey dynamics (Figure S16). This suggests that while a saturating prey-dependent functional response may well-describe the dynamics of the predators, it may not fully capture the prey population dynamics in this system. If the functional responses for prey and predator are assumed to be symmetric, then the Rosenzweig-MacArthur model fail to describe the dynamics.
- **Hassell-Verley model**. The Hassell-Varley model does not perform well for either prey or predator dynamics. This indicates that a non-saturating ratio-dependent functional response is not appropriate for describing the dynamics observed in this case study.
- **Lotka-Volterra model**. The Lotka-Volterra model without self-regulation performs well for most experimental replicates except for replicates 8, 9, and 10 for the prey dynamics. In these three replicates, the environmental conditions were not constant, in contrast to the other seven replicates. This may explain why the Lotka-Volterra dynamics with self-regulation explain the prey dynamics better in these three cases.

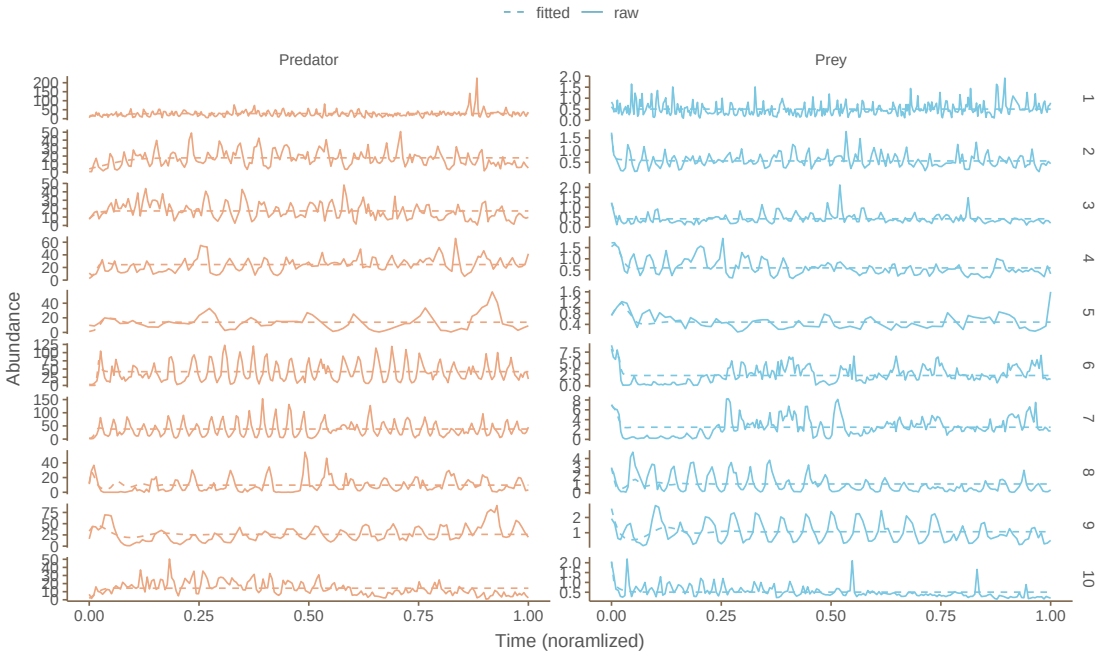
In conclusion, it reinforces our original conclusion that the invalidation of the ratio-dependent model in the case study.

Supplementary Note 8 Comparison with traditional methods

Here we show the results with classic valiodation methods.

8.1 Regression on inferred deriatives

Figure S18 below illustrates the results:

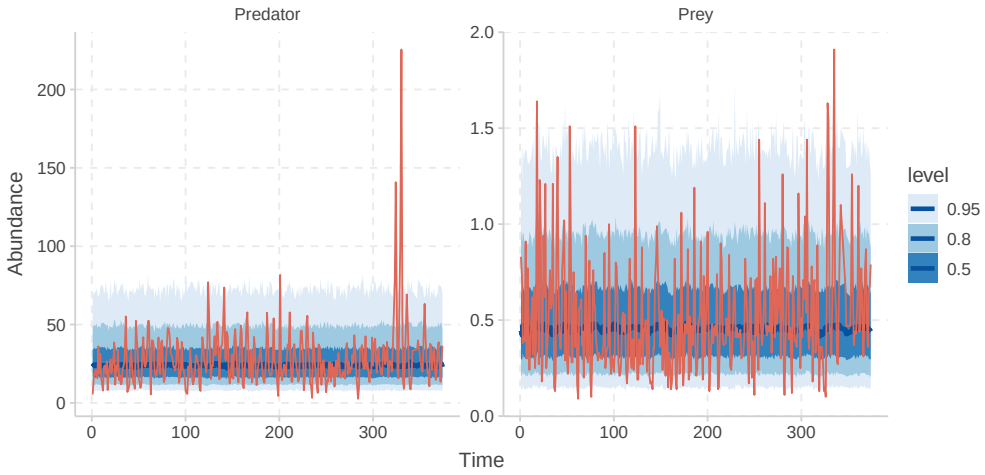


Supplementary Figure S18: Time series of prey and predator abundances using regression on inferred derivatives. The thick lines show the observed dynamics, while the dashed lines show the best fitted models. The models capture the mean trends but not the oscillatory behavior. Columns represent different species (prey and predator), and rows represent the 10 replicates.

Despite the optimization and smoothing techniques, the regression-based methods captured only the mean trends and failed to reproduce the oscillatory behavior observed in the empirical data. We have also explored Lasso and Ridge regression to enhance model performance by mitigating overfitting and handling noise, but they did not provide any improvement.

8.2 Bayesian nonlinear ODE modeling

Given the computationally intensive nature of this approach, we have only used it for the main expeirment (replicate 1) in Blasius *et al.* (2020). We expect similar results for other replicates, although further investigation is needed to confirm this. Figure S19 belows shows the posterior predictive checks of the Bayesian model:



Supplementary Figure S19: Bayesian nonlinear ODE model fit to prey and predator abundances. The observed values are shown in orange, while the black line represents the posterior mean prediction with 50%, 80%, and 95% credible intervals (shaded areas). The model captures the main trends but not the oscillatory behavior

The Bayesian model effectively captured the main trends in the data but, similar to the regression approach, failed to reproduce the observed oscillations.

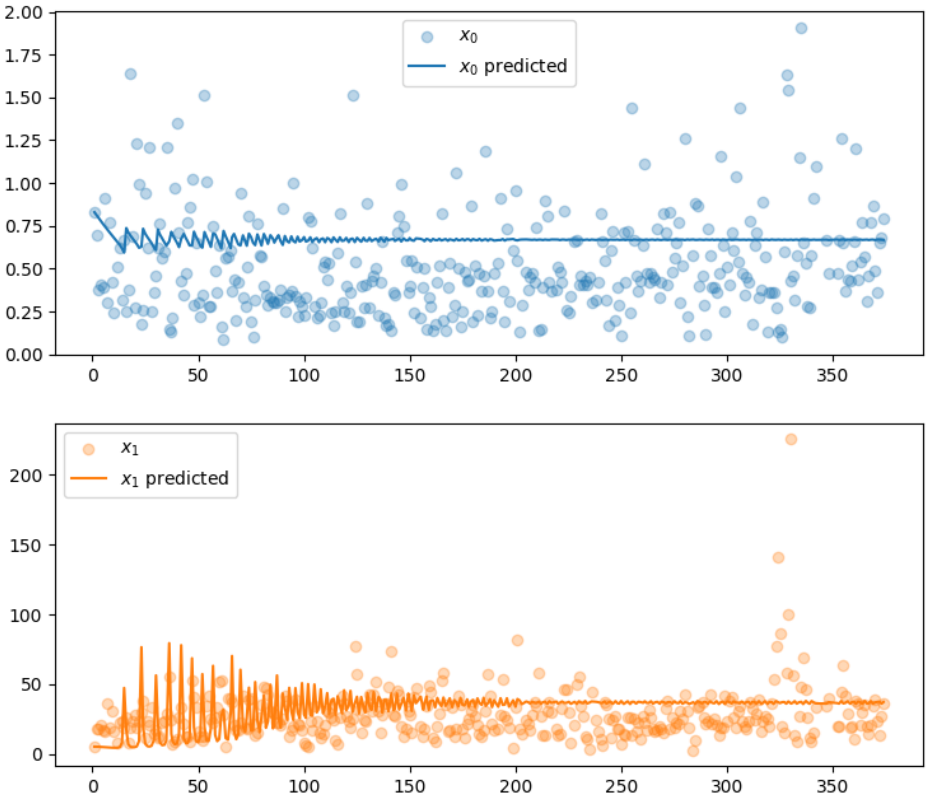
8.3 Symbolic regression with deep learning

For the main experiment (replicate 1) in Blasius *et al.* (2020), the best inferred equations read:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -0.2710N + 0.0048P \tag{S24}$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = 14.2024(1 + 0.0034P)^2 \sin(-1.0600 + 21.8434N) \tag{S25}$$

These equations lack ecological interpretability and do not align with any known predator-prey model. The model's fit to the data was also poor, as illustrated in Figure S20 below:



Supplementary Figure S20: Symbolic regression results for prey (x_0) and predator (x_1) abundances. The model fails to capture the observed dynamics.

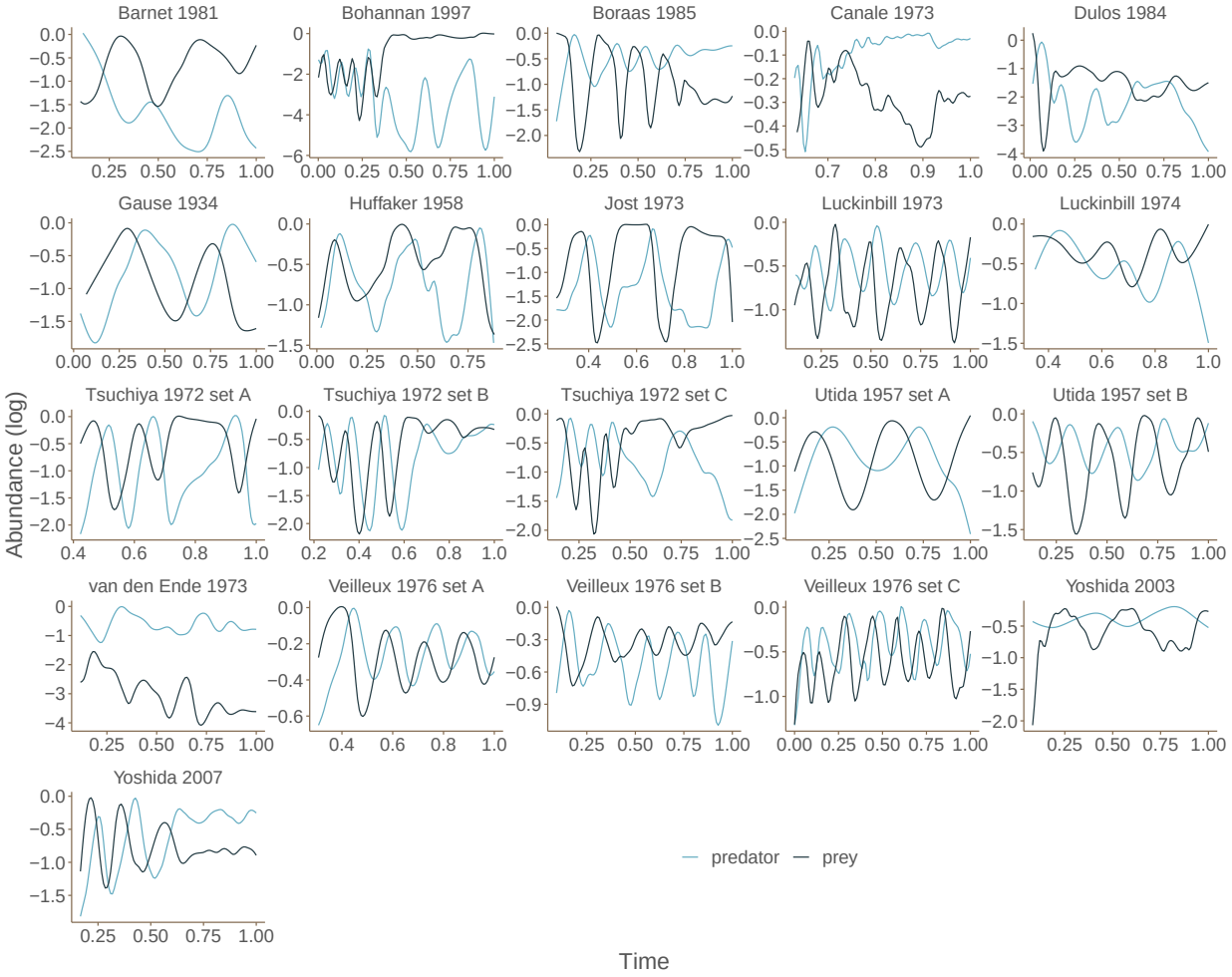
Across all ten replicates, the R^2 values ranged from -0.54 (the minus sign is not a typo) to 0.06, indicating inadequate model performance. This suggests that the method, despite its sophistication, struggled with ecological data and produced non-generalizable results.

Supplementary Note 9 Dissecting ecological and evolutionary processes

We analyzed 21 two-species consumer–resource time series (Barnet *et al.*, 1981; Bohannan & Lenski, 1997; Dulos & Marchand, 1984; Canale *et al.*, 1973; Jost *et al.*, 1973; Van den Ende, 1973; Tsuchiya *et al.*, 1972; Gause, 1934; Luckinbill, 1973, 1974; Veilleux, 1976; Boraas, 1980; Yoshida *et al.*, 2003, 2007; Huffaker *et al.*, 1958; Utida, 1957). These datasets were gathered by Hiltunen *et al.* (2014).

9.1 Visualizing raw data

Figure S21 shows the time series in the system.



Supplementary Figure S21: Time series of prey and predator abundances.

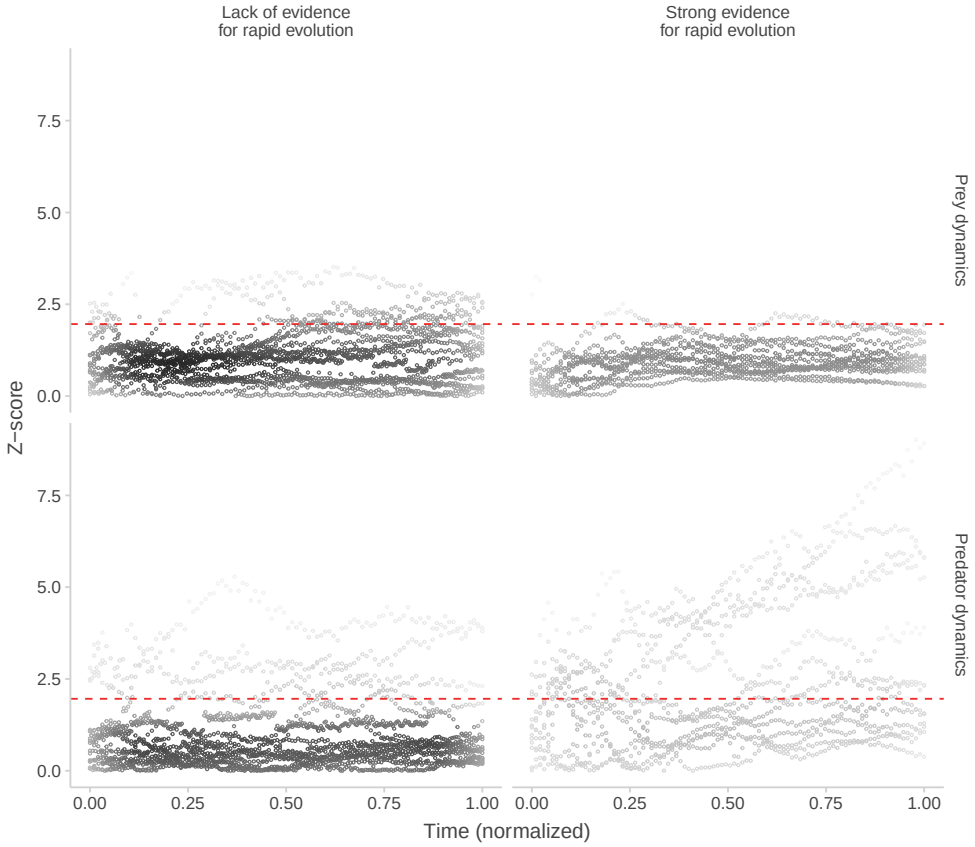
9.2 Stationarity test

	Non-stationary	Stationary
w/ rapid evolution		
predator	0.23	0.77
prey	0.23	0.77
w/o rapid evolution		
predator	0.38	0.62
prey	0.12	0.88

Table Supplementary Table 2: Summary of proportions of time series that are non stationary and stationary.

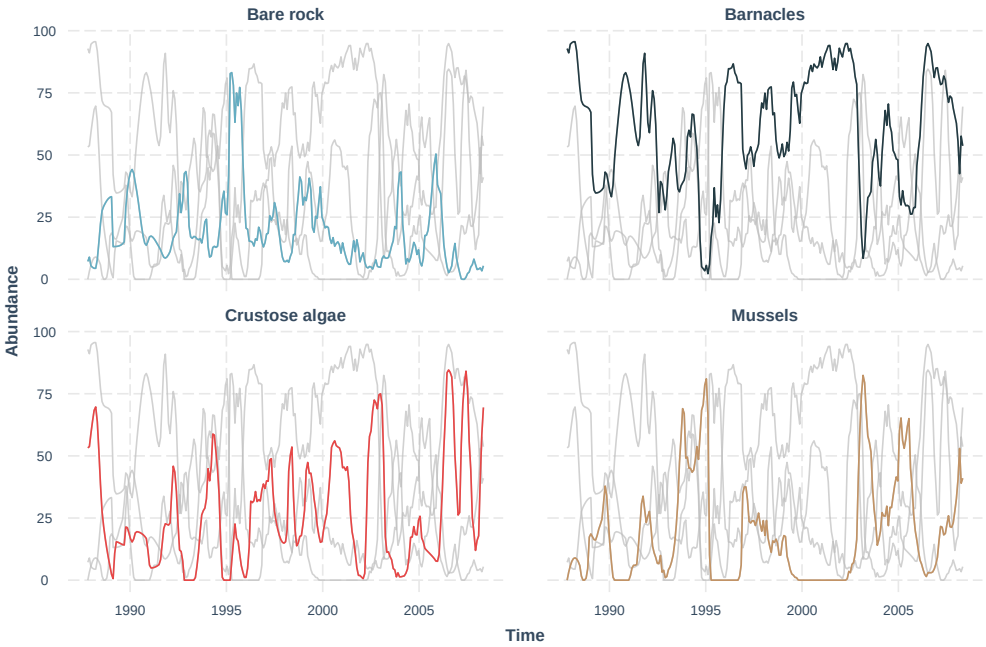
9.3 Moving-window analysis

To understand the effect of time series length, we performed a moving-window analysis. Each window starts at the origin of the time series. We progressively increased the length of these windows, beginning with a small initial window and expanding until the full dataset was encompassed. This approach allowed us to evaluate the performance of our covariance criterion across different time spans within the dataset.



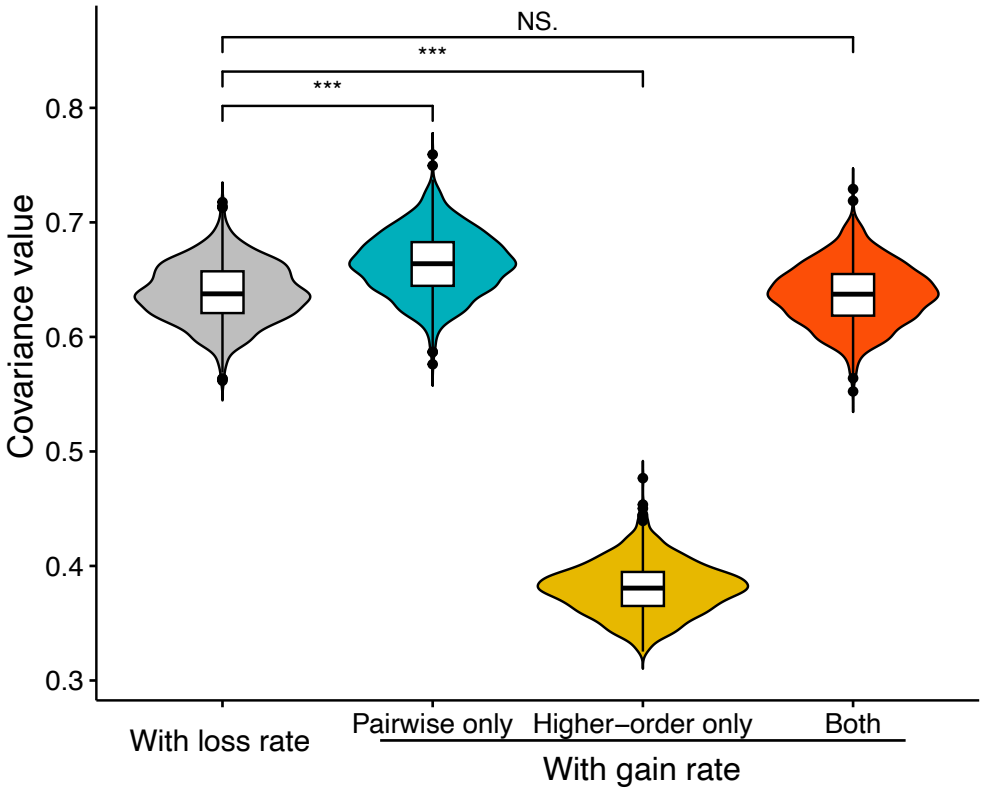
Supplementary Figure S22: Statistical test of covariance criteria.

Supplementary Note 10 Detecting signals of higher-order interactions



Supplementary Figure S23: Time series.

The dynamics for all species are stationary, with p value < 0.01 with augmented Dickey-Fuller test.



Supplementary Figure S24: **Detecting signals of higher-order interactions**. We apply the covariance criteria to study the presence of higher-order interactions in a rocky intertidal community (Benincà *et al.*, 2015). Specifically, whether crustose algae and barnacles interact with mussels exclusively through pairwise interactions or whether a higher-order interaction is present. Three models are evaluated: pairwise interactions only (blue), higher-order interaction only (yellow), and a combination of both (red). The x-axis denotes different covariance types, while the y-axis shows normalized covariance values derived from 1000 bootstrapping replicates. The gray boxplot represents the covariance with loss rate, assumed constant across models. The higher-order only model (yellow) exhibits significant mismatch in covariance values, indicating its inadequacy. The pairwise interaction model (blue) shows closer alignment but still differs statistically from the loss covariance. Remarkably, the model incorporating both pairwise and higher-order interactions (red) perfectly matches the loss covariance. These findings strongly suggest that both pairwise and higher-order interactions between crustose algae, barnacles, and mussels play a significant role in influencing mussel dynamics within this community.

References

Abrams, P. A. (1982). Functional responses of optimal foragers. *The American Naturalist*, 120, 382–390.

Abrams, P. A. (1990). The effects of adaptive behavior on the type-2 functional response. *Ecology*, 71, 877–885.

Aldebert, C., Nerini, D., Gauduchon, M. & Poggiale, J. (2016a). Does structural sensitivity alter complexity–stability relationships? *Ecological complexity*, 28, 104–112.

Aldebert, C., Nerini, D., Gauduchon, M. & Poggiale, J. (2016b). Structural sensitivity and resilience in a predator–prey model with density-dependent mortality. *Ecological complexity*, 28, 163–173.

Andrews, J. F. (1968). A mathematical model for the continuous culture of microorganisms utilizing inhibitory substrates. *Biotechnology and bioengineering*, 10, 707–723.

Arditi, R. & Akçakaya, H. (1990). Underestimation of mutual interference of predators. *Oecologia*, 83, 358–361.

Arditi, R. & Ginzburg, L. R. (1989). Coupling in predator-prey dynamics: ratio-dependence. *Journal of Theoretical Biology*, 139, 311–326.

Barbier, M., Wojcik, L. & Loreau, M. (2021). A macro-ecological approach to predation density-dependence. *Oikos*, 130, 553–570.

Barnet, Y. M., Daft, M. & Stewart, W. (1981). Cyanobacteria-cyanophage interactions in continuous culture. *Journal of Applied Bacteriology*, 51, 541–552.

Beddington, J. R. (1975). Mutual interference between parasites or predators and its effect on searching efficiency. *The Journal of Animal Ecology*, 331–340.

Benincà, E., Ballantine, B., Ellner, S. P. & Huisman, J. (2015). Species fluctuations sustained by a cyclic succession at the edge of chaos. *Proceedings of the National Academy of Sciences*, 112, 6389–6394.

Blasius, B., Rudolf, L., Weithoff, G., Gaedke, U. & Fussmann, G. F. (2020). Long-term cyclic persistence in an experimental predator–prey system. *Nature*, 577, 226–230.

Bohannan, B. J. & Lenski, R. E. (1997). Effect of resource enrichment on a chemostat community of bacteria and bacteriophage. *Ecology*, 78, 2303–2315.

Boraas, M. E. (1980). A chemostat system for the study of rotifer-algal-nitrate interactions. *Evolution and ecology of zooplankton communities*.

Canale, R. P., Lustig, T., Kehrberger, P. M. & Salo, J. (1973). Experimental and mathematical modeling studies of protozoan predation on bacteria. *Biotechnology and Bioengineering*, 15, 707–728.

Cosner, C., DeAngelis, D. L., Ault, J. S. & Olson, D. B. (1999). Effects of spatial grouping on the functional response of predators. *Theoretical population biology*, 56, 65–75.

Crowley, P. H. & Martin, E. K. (1989). Functional responses and interference within and between year classes of a dragonfly population. *Journal of the North American Benthological Society*, 8, 211–221.

DL, D. (1975). A model for trophic interaction. *Ecology*, 56, 881–892.

Dulos, E. & Marchand, A. (1984). Oscillations of the population densities of the bacterial prey-predator couple escherichia coli-bdellovibrio bacteriovorus: experimental study and theoretical model. In: *Annals of the Institut Pasteur/Microbiology*, vol. 135. Elsevier.

Fujii, K., Holling, C. & Mace, P. (1986). A simple generalized model of attack by predators and parasites. *Ecological Research*, 1, 141–156.

Gause, G. F. (1934). Experimental analysis of vito volterra's mathematical theory of the struggle for existence. *Science*, 79, 16–17.

Gilpin, M. E. & Ayala, F. J. (1973). Global models of growth and competition. *Proceedings of the National Academy of Sciences*, 70, 3590–3593.

Gutierrez, A., Baumgärtner, J. & Summers, C. (1984). Multitrophic models of predator–prey energetics: I. age-specific energetics models—pea aphid acyrthosiphon pisum (homoptera: Aphididae) as an example1. *The Canadian Entomologist*, 116, 923–932.

Hassell, M., Lawton, J. & Beddington, J. (1977). Sigmoid functional responses by invertebrate predators and parasitoids. *The Journal of Animal Ecology*, 249–262.

Hassell, M. & Varley, G. (1969). New inductive population model for insect parasites and its bearing on biological control. *Nature*, 223, 1133—1137.

Hatton, I. A., Mazzarisi, O., Altieri, A. & Smerlak, M. (2024). Diversity begets stability: Sublinear growth and competitive coexistence across ecosystems. *Science*, 383, eadg8488.

Hilfinger, A., Norman, T. M., Vinnicombe, G. & Paulsson, J. (2016). Constraints on fluctuations in sparsely characterized biological systems. *Physical Review Letters*, 116, 058101.

Hiltunen, T., Hairston Jr, N. G., Hooker, G., Jones, L. E. & Ellner, S. P. (2014). A newly discovered role of evolution in previously published consumer–resource dynamics. *Ecology Letters*, 17, 915–923.

Holling, C. S. (1959). Some characteristics of simple types of predation and parasitism1. *The canadian entomologist*, 91, 385–398.

Holling, C. S. (1965). The functional response of predators to prey density and its role in mimicry and population regulation. *The Memoirs of the Entomological Society of Canada*, 97, 5–60.

Huffaker, C. *et al.* (1958). Experimental studies on predation: dispersion factors and predator-prey oscillations. *Hilgardia*, 27, 343–383.

Ivlev, V. S. (1955). *Experimental Ecology of the Feeding of Fishes*. Yale University Press.

Jassby, A. D. & Platt, T. (1976). Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. *Limnology and oceanography*, 21, 540–547.

Jeschke, J. M., Kopp, M. & Tollrian, R. (2002). Predator functional responses: discriminating between handling and digesting prey. *Ecological monographs*, 72, 95–112.

Jost, J., Drake, J., Fredrickson, A. & Tsuchiya, H. (1973). Interactions of tetrahymena pyriformis, escherichia coli, azotobacter vinelandii, and glucose in a minimal medium. *Journal of Bacteriology*, 113, 834–840.

Kratina, P., Vos, M., Bateman, A. & Anholt, B. R. (2009). Functional responses modified by predator density. *Oecologia*, 159, 425–433.

Lotka, A. (1925). Elements of physical biology. *Williams and Wilkins*.

Luckinbill, L. S. (1973). Coexistence in laboratory populations of paramecium aurelia and its predator didinium nasutum. *Ecology*, 54, 1320–1327.

Luckinbill, L. S. (1974). The effects of space and enrichment on a predator-prey system. *Ecology*, 55, 1142–1147.

Mazzarisi, O. & Smerlak, M. (2024). Complexity-stability relationships in competitive disordered dynamical systems. *Physical Review E*, 110, 054403.

Michaelis, L., Menten, M. L. *et al.* (1913). Die kinetik der invertinwirkung. *Biochem. z*, 49, 352.

Novak, M. & Stouffer, D. B. (2021). Geometric complexity and the information-theoretic comparison of functional-response models. *Frontiers in Ecology and Evolution*, 9, 740362.

Okuyama, T. & Ruyle, R. L. (2011). Solutions for functional response experiments. *Acta oecologica*, 37, 512–516.

Pennekamp, F., Iles, A. C., Garland, J., Brennan, G., Brose, U., Gaedke, U., Jacob, U., Kratina, P., Matthews, B., Munch, S. *et al.* (2019). The intrinsic predictability of ecological time series and its potential to guide forecasting. *Ecological Monographs*, 89, e01359.

Pimm, S. L. (1982). *Food webs*. Springer.

Real, L. A. (1977). The kinetics of functional response. *The American Naturalist*, 111, 289–300.

Rogers, T. L., Johnson, B. J. & Munch, S. B. (2022). Chaos is not rare in natural ecosystems. *Nature Ecology & Evolution*, 6, 1105–1111.

Rosenzweig, M. L. (1971). Paradox of enrichment: destabilization of exploitation ecosystems in ecological time. *Science*, 171, 385–387.

Ruxton, G., Gurney, W. & De Roos, A. (1992). Interference and generation cycles. *Theoretical Population Biology*, 42, 235–253.

Schenk, D., BERSIER, L.-F. & Bacher, S. (2005). An experimental test of the nature of predation: neither prey-nor ratio-dependent. *Journal of Animal Ecology*, 74, 86–91.

Sokol, W. & Howell, J. (1981). Kinetics of phenol oxidation by washed cells. *Biotechnology and Bioengineering*, 23, 2039–2049.

Stouffer, D. B. & Novak, M. (2021). Hidden layers of density dependence in consumer feeding rates. *Ecology Letters*, 24, 520–532.

Sutherland, W. J. (1983). Aggregation and theideal free'distribution. *The Journal of Animal Ecology*, 821–828.

Thomas, W. R., Pomerantz, M. J. & Gilpin, M. E. (1980). Chaos, asymmetric growth and group selection for dynamical stability. *Ecology*, 61, 1312–1320.

Tostowaryk, W. (1972). The effect of prey defense on the functional response of podisus modestus (hemiptera: Pentatomidae) to densities of the sawflies neodiprion swainei and n. pratti banksianae (hymenoptera: Neodiprionidae). *The Canadian Entomologist*, 104, 61–69.

Tsuchiya, H., Drake, J., Jost, J. & Fredrickson, A. (1972). Predator-prey interactions of dictyostelium discoideum and escherichia coli in continuous culture. *Journal of Bacteriology*, 110, 1147–1153.

Tyutyunov, Y., Titova, L. & Arditi, R. (2008). Predator interference emerging from trophotaxis in predator–prey systems: an individual-based approach. *ecological complexity*, 5, 48–58.

Utida, S. (1957). Population fluctuation, an experimental and theoretical approach. In: *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 22. Cold Spring Harbor Laboratory Press.

Van den Ende, P. (1973). Predator-prey interactions in continuous culture. *Science*, 181, 562–564.

Veilleux, B. (1976). *The analysis of a predatory interaction between Didinium and Paramecium.* Msc thesis, University of Alberta, Canada.

Volterra, V. (1927). Fluctuations in the abundance of a species considered mathematically. *Nature*, 119, 12–13.

Watt, K. (1959). A mathematical model for the effect of densities of attacked and attacking species on the number attacked. *The Canadian Entomologist*, 91, 129–144.

Yoshida, T., Ellner, S. P., Jones, L. E., Bohannan, B. J. M., Lenski, R. E. & Hairston Jr, N. G. (2007). Cryptic population dynamics: rapid evolution masks trophic interactions. *PLoS Biology*, 5, e235.

Yoshida, T., Jones, L. E., Ellner, S. P., Fussmann, G. F. & Hairston, N. G. (2003). Rapid evolution drives ecological dynamics in a predator–prey system. *Nature*, 424, 303–306.