

# Rigorous validation of ecological models against empirical time series

Received: 23 September 2024

Accepted: 27 August 2025

Published online: 27 October 2025

 Check for updates

Chuliang Song   & Jonathan M. Levine   

The complexity of ecosystems poses a formidable challenge for validating ecological models. The prevailing inability to falsify models has resulted in an accumulation of models but not an accumulation of confidence. Here we introduce an approach rooted in queueing theory, termed the covariance criteria, that establishes a rigorous test for model validity based on covariance relationships between observable quantities. These criteria set a high bar for models to pass by specifying necessary conditions that must hold regardless of unobserved factors. We test our approach using observed time series data on three long-standing challenges in ecological theory: resolving competing models of predator-prey functional responses, disentangling ecological and evolutionary dynamics in systems with rapid evolution and detecting the often-elusive influence of higher-order species interactions. Across these diverse case studies, the covariance criteria consistently rule out inadequate models, while building confidence in those that provide strategically useful approximations. The covariance criteria approach is mathematically rigorous and computationally efficient, making it applicable to existing data and models.

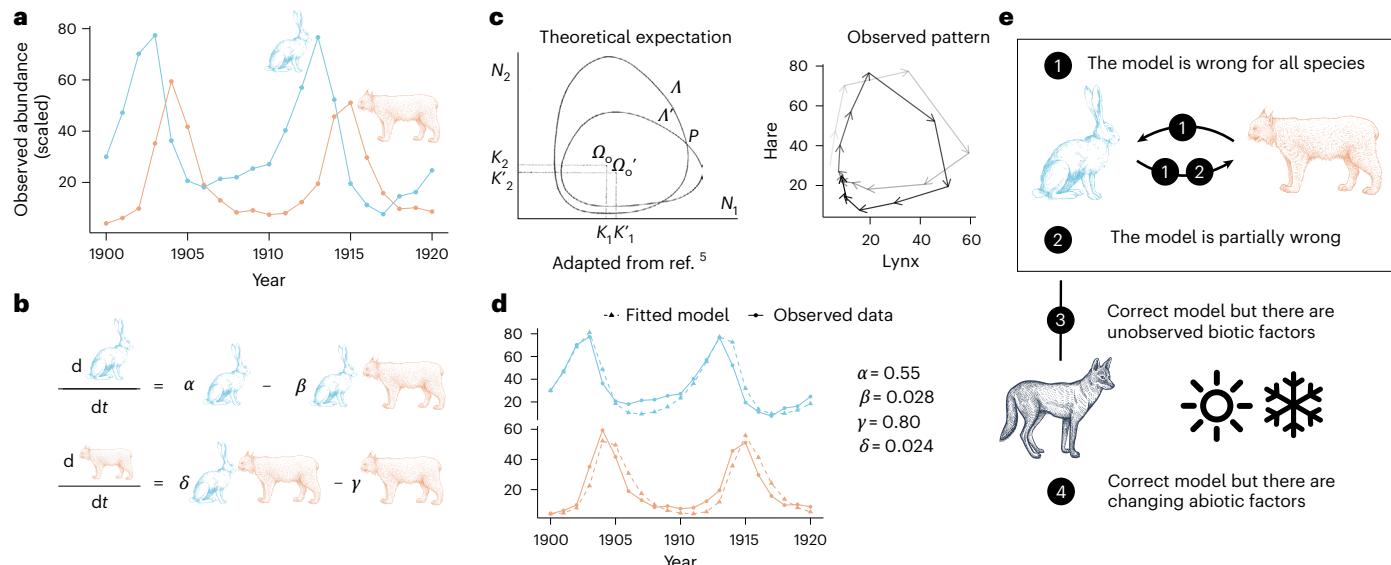
Population abundance is the ever-present variable in the equation of life on Earth. To decipher the drivers behind the fluctuations of population abundance, ecologists construct mathematical models—simplified representations that capture the key dynamics of an ecosystem while making judicious sacrifices of the full complexity of Nature. This synergy between data and modelling forms the foundation of contemporary ecology<sup>1,2</sup>. Yet, this endeavour faces a fundamental challenge: how can we confidently adjudicate which models provide useful approximations of Nature, and which are oversimplified caricatures? The stark reality is that even for predator-prey interactions, there exist more than 40 distinct models of how predator feeding rate depends on prey abundance (reviewed in ref. 3). This plethora of alternatives stems from the prevailing inability, using conventional practices, to decisively validate some models and invalidate others against empirical data.

To illustrate the limitations of current approaches, consider a textbook example of the coupled population dynamics of snowshoe

hares and Canadian lynx in boreal forests<sup>4</sup> (Fig. 1a). The dynamics is classically modelled using the Lotka–Volterra predator–prey model (Fig. 1b). A common validation approach is to compare the qualitative behaviours between data and model prediction (Fig. 1c). In this example, the Lotka–Volterra model predicts coupled cycles of predator and prey abundances with a fixed amplitude and period length—the hallmark of the ‘predation cycle’ in graphical predator–prey theory<sup>5,6</sup>. The qualitative resemblance of the data to these predicted cycles provides some confidence in the validity of the model. Another common approach is fitting models to data and assessing goodness-of-fit or forecasting power (Fig. 1d). For this example, the predator–prey dynamics can be approximated with a given set of parameters in Lotka–Volterra dynamics, providing support for the proposed model. These two approaches represent the mainstream for validating models against time series data.

But what if the model and data diverge? Does that mean the model is invalidated? Returning to the hare–lynx example, over longer

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA. <sup>2</sup>Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA.  e-mail: [chuliang.song@ucla.edu](mailto:chuliang.song@ucla.edu)



timescales than shown in Fig. 1, the observed population cycles diverge from the classic Lotka–Volterra predictions: in fact, the data suggest a ‘reversed cycle’, implying the nonsensical result that hares eat lynx<sup>7</sup>. However, interpreting such discrepancies is fraught with ambiguity. One possibility is that these deviations in qualitative patterns invalidate the core assumptions of the Lotka–Volterra framework, demanding an alternative theoretical model for the entire predator–prey system (case 1 in Fig. 1e)<sup>8</sup>. Alternatively, the model could accurately describe the dynamics of one species (for example, lynx) while failing for the other (for example, hares) due to missing factors specific to that species (case 2 in Fig. 1e)<sup>9</sup>. In contrast, it is also possible that the model correctly describes the lynx–hare interaction, but fails to include all the other variables driving the observed dynamics. These could be other uncontrolled biotic factors (case 3 in Fig. 1e), such as the hare–vegetation interaction<sup>10</sup> or disease epidemics. Or it could be that there are other uncontrolled abiotic factors (case 4 in Fig. 1e), such as environmental fluctuations altering species parameters over time<sup>11–13</sup>. In sum, current practices make it challenging to judge whether the model is truly valid, partially valid or simply wrong.

This long-standing challenge of model validation has plagued ecology, leaving the true scope of even classic models such as the Lotka–Volterra formulation unresolved. Here we address this fundamental problem in ecological modelling by introducing a method originally developed by biophysicists<sup>14</sup>. In essence, this method uncovers the (mostly unique) inherent structure of temporal covariance between model elements, a constraint that remains invariant regardless of unknown ecological factors. By leveraging this inherent constraint, we can make strong statements about model validity. In the following sections, we first introduce the theoretical foundations of this approach. We then demonstrate its discriminatory power by applying it to three key problems in ecology: resolving debates on the functional form of predator–prey interactions, disentangling the interplay of ecology and rapid evolution, and detecting signals of higher-order species interactions. Through these case studies, we illustrate how this rigorous test of model validity can decisively invalidate flawed models, build

confidence in those that provide useful approximations and guide the development of more robust ecological theory.

## Covariance criteria for model (in)validation

In this section, we introduce the theoretical framework for the covariance criteria approach and demonstrate its application to ecological models and data. We start by presenting the core concepts and mathematical foundations. We then illustrate how to apply the framework to a simple worked-out example with statistical methods. We then discuss the advantages of this approach over current model validation practices, as well as the caveats.

### General theoretical framework

The fluctuations in population abundances that we observe in Nature arise from a fundamental imbalance between two opposing forces: the gain rate, which encompasses processes that increase population size (for example, births, immigration and mutualism) and the loss rate, which includes processes that decrease it (for example, deaths, emigration and competition). In general, ecological models describing the dynamics of population abundance can be partitioned in the following form:

$$\frac{dx}{dt} = R_+ - R_- + \xi, \quad (1)$$

Population abundance

↓

$\frac{dx}{dt}$

$R_+$  Gain rate

$R_-$  Loss rate

Stochastic noise

where  $R_+$  is the gain rate,  $R_-$  is the loss rate and  $\xi$  is the stochastic noise. The gain rate  $R_+$  and loss rate  $R_-$  can be complex functions of both biotic (for example, interactions with other species) and abiotic (for example, environmental) factors. This partitioning is not arbitrary; it follows

directly from the mathematical formulation of a model, with  $R_+$  comprising the sum of all positive terms and  $R_-$  the sum of the absolute values of all negative terms. The specific structure of a model therefore represents a distinct, testable hypothesis about the underlying gain and loss processes (Methods). Any meaningful ecological model must include this combination of gain and loss; otherwise, the model would predict either indefinite growth or inevitable extinction.

We can illustrate this partitioning using the predator dynamics from the Lotka–Volterra model: where  $x$  denotes the predator and  $y$  denotes the prey.

When gain and loss rates are perfectly counterbalanced, the population maintains a steady state or equilibrium. However, such perfect balance is seldom encountered in the real world. Instead, we witness periods where gains outweigh losses, leading to population growth, interspersed with periods where losses dominate, causing population decline. This interplay between gains and losses generates the dynamic fluctuations in abundance that characterize most natural populations.

While it is intuitive that population abundance ( $x$ ) fluctuates based on the balance of gains ( $R_+$ ) and losses ( $R_-$ ), mathematically formalizing this link presents a challenge. The core difficulty stems from the fact that the gain–loss imbalance ( $R_+ - R_-$ ) directly dictates the instantaneous rate of change in abundance ( $\frac{dx}{dt}$ ), not the absolute abundance ( $x$ ) at any given time. In plain words, the fundamental equation of population dynamics (equation (1)) connects the processes driving change (gain and loss rates) to the speed at which abundance changes, rather than directly to the current amount of abundance itself. This makes it less straightforward to quantify how the underlying gain and loss rates relate to the population abundance as it varies over time. To overcome this and directly assess the relationship between the gain/loss processes and the raw population abundance, we use covariance. Covariance provides a statistical measure of how two variables—in this case, a rate (such as  $R_+$ ) and the absolute abundance ( $x$ )—tend to fluctuate together over the observed time series. For instance, if gain rates consistently tend to be higher when population abundance is high, these two variables exhibit a positive covariance. Crucially, for populations exhibiting bounded fluctuations, the loss rate must also, on average, increase when abundance is high to counterbalance the increased gain. This necessary regulatory feedback implies that the loss rate and population abundance should also covary positively. By examining these covariance relationships, specifically how the gain and loss rates covary with abundance, we can derive quantitative criteria to test the structural assumptions of ecological models.

Formally, this constraint, while grounded in sophisticated mathematics, is captured in a surprisingly simple equality<sup>14</sup>: where Cov denotes covariance and  $\langle \rangle$  denotes mean (Fig. 2a). In words, the equality (equation (3)) essentially states that the normalized covariance between gain rate and abundance is mirrored by the normalized covariance between loss rate and abundance. Note that this normalized covariance is distinct from the correlation coefficient, which is normalized by standard deviations rather than means.

$$\frac{\text{Cov}(R_+, R)}{\langle R_+ \rangle \langle x \rangle} = \frac{\text{Cov}(R_-, R)}{\langle R_- \rangle \langle x \rangle} \quad (3)$$

Mathematically, this equality is known as the second-order moment equation derived from Little's law in queuing theory<sup>15,16</sup>. It

is a continuous-time formulation of the more general constraints<sup>14</sup>. Specifically, our equation (3) can be derived from their equation (6) by setting the two state variables to be the same ( $i=j$ ), which simplifies the cross-correlation to an autocorrelation, and then taking the continuous-time limit where the discrete time step vanishes. We adopt a continuous framework because population dynamics in ecology are commonly modelled with differential equations and empirical data often represent continuous variables such as biomass or density (which do not have a direct discrete interpretation).

## An illustrated worked-out example

This covariance structure serves as a simple test to validate or invalidate a model. If the model accurately reflects the observed ecological dynamics, the equality (equation (3)) will hold and the model passes the test. If the data and constraint do not match, the equality will not hold and the model is falsified. We call this approach the covariance criteria.

We next illustrate how the covariance criteria works in practice using the predator dynamics from the Lotka–Volterra model (equation (2)). The model defines the gain rate as proportional to the product of prey and predator abundances (representing successful predation leading to reproduction), while the loss rate is proportional to predator abundance alone (representing mortality). Applying the general covariance constraint (equation (3)) to this model, we get: Since the mean and covariance operators are linear, we can simplify this further by cancelling out the constant parameters  $\delta$  and  $\gamma$ :

$$\frac{\text{Cov}(\delta yx, x)}{\langle \delta yx \rangle \langle x \rangle} = \frac{\text{Cov}(yx, x)}{\langle yx \rangle \langle x \rangle}. \quad (4)$$

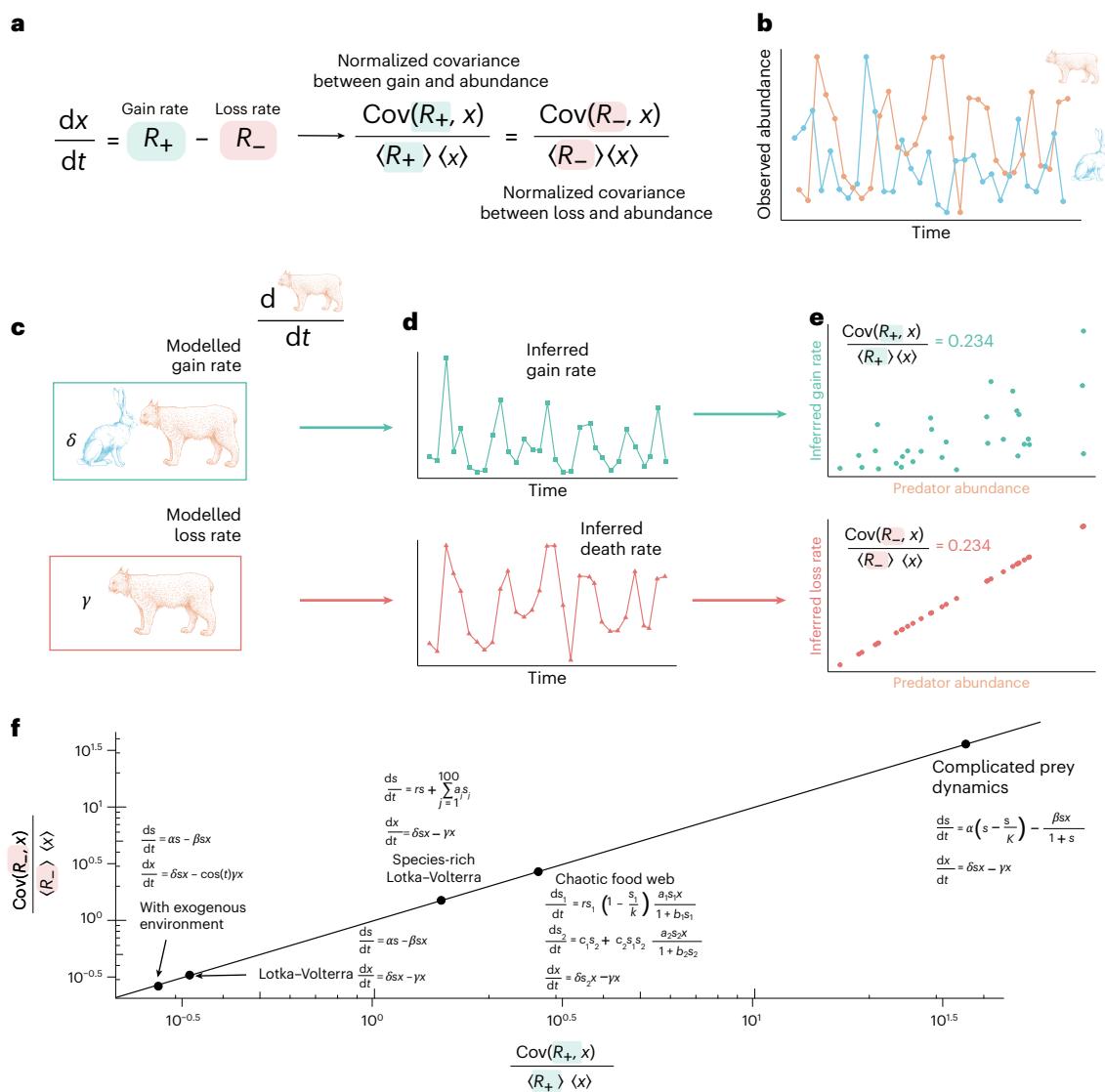
$$\frac{\text{cov}(yx, x)}{\langle yx \rangle \langle x \rangle} = \frac{\text{cov}(x, x)}{\langle x \rangle \langle x \rangle}. \quad (5)$$

This simplification is powerful: it allows us to test the structural form of the Lotka–Volterra model using only the observed time series data for predator ( $x$ ) and prey ( $y$ ) abundances, without needing to estimate the parameter values  $\delta$  and  $y$ . In such cases, the covariance criterion provides a non-parametric test of the structure of the model.

We next test this model using a subset of data from a planktonic predator–prey system<sup>17</sup> (Fig. 2b). The data include time series measurements of both predator and prey abundance. Although we cannot directly observe the gain and loss rates in the data, the model allows us to infer them at each time point on the basis of the observed abundance data (Fig. 2c). For instance, the gain rate is inferred as the product of prey and predator abundance at each time point (recall that the parameter in front of this product cancels out in equation (5)).

The covariance criteria then examine how these inferred gain and loss rates covary with the observed predator abundance (Fig. 2c). The scatter plot of inferred gain rate against predator abundance shows more scatter compared with the plot of inferred loss rate against predator abundance. This is expected because the predator's gain rate depends on both prey and predator abundance, while the loss rate depends solely on predator abundance. However, it is important to remember that the covariance criteria rely on the calculated covariance values to provide the quantitative measure of the relationship between the rates and predator abundance, not the visual spread of the scatter plots. In this specific example, both normalized covariances turn out to be 0.234. This suggests that the Lotka–Volterra model, in this case, aligns with the covariance criterion and adequately explains the observed predator dynamics.

While the Lotka–Volterra example illustrates the use of the covariance criteria when the model contains simple gain and loss terms, the approach is equally valid for models with more complex relationships between gain or loss terms and species abundance. As an example, Supplementary Note 1 shows that the covariance criteria can be used



**Fig. 2 | Covariance criteria for model validation.** **a**, The general principle of covariance criteria (theoretical framework): if a model accurately captures the underlying processes, the normalized covariance between the gain rate and abundance should equal the normalized covariance between the loss rate and abundance. **b–e**, The workflow of this approach to assess whether the Lotka–Volterra model can describe predator dynamics in a specific system: a subset of the predator–prey observed data<sup>17</sup> (**b**); model for predator demonstrating how the Lotka–Volterra model partitions predator dynamics into gain rate (terms causing abundance increase) and loss rate (terms causing abundance decrease) (**c**); integrating the modelled gain and loss rates with the observed data to infer the empirical gain and loss rates across time (**d**); and examining covariance structure by calculating the normalized covariances between the

inferred gain/loss rates and predator abundance (**e**). If these covariances are equivalent, it strengthens confidence in the validity of the model. Conversely, significant discrepancies indicate the inadequacy of the Lotka–Volterra model in describing the predator's dynamics. **f**, Covariance structure holds irrespective of the unassumed, that is emphasizing the universality of the covariance criteria (Methods). This means that the criteria hold true even if the model does not explicitly include all factors influencing the system. To illustrate this, simulations are used where the predator follows the Lotka–Volterra model, but the rest of the ecological community can exhibit arbitrarily complex dynamics. These simulations serve purely as an illustrative aid, as the core strength of the method lies in its mathematical rigour. Illustrations: mashikomo/stock.adobe.com.

to evaluate the MacArthur–Rosenzweig predator–prey model with a type II functional response.

#### Discriminatory power of the covariance criteria

The covariance criteria work under remarkably broad conditions, thanks to the generality of Little's Law in queuing theory<sup>15,16</sup>. It applies rigorously to stationary systems, where long-term statistical patterns remain constant over time, regardless of whether they follow typical Markovian dynamics (where the future depends only on the present, as most ecological models do) or more complex non-Markovian dynamics (where the history of the ecosystem influences its future, as in the presence of time delays). Moreover, the criteria also hold for some

non-stationary dynamics, such as cyclo-stationary systems (where statistical patterns repeat predictably, as with seasons).

One might assume that such broad applicability renders the criteria a mere abstract principle with limited practical utility. Surprisingly, as we show here, it imposes a stringent test for models to pass. When the biophysicists who pioneered this method applied it to gene expression data, nearly all published models failed to meet the criteria<sup>18</sup>. This high bar means that when a model does pass the test, we can have strong confidence in its validity as a useful approximation of the true dynamics of the system.

The stringency of the covariance criteria arises because different ecological model structures typically yield distinct covariance

structures—their own unique quantitative ‘fingerprints’. To assess this discriminatory power systematically, we applied the covariance criteria framework to simulations of all 40 predator functional response models reviewed by ref. 3. While many of these models produce qualitatively similar cyclic dynamics, our analysis revealed that each possesses a unique, distinguishable covariance structure (Supplementary Note 2). This demonstrates the potential of the covariance criteria to differentiate between competing models that might otherwise be indistinguishable based solely on visual inspection or standard goodness-of-fit metrics.

Leveraging this discriminatory power and the underlying generality of the model allows the covariance criteria to interrogate models with greater precision than traditional approaches, overcoming common ambiguities in validation (Fig. 1e). First, the criteria directly test the dynamics of individual species, eliminating the need to know whether the model is correct for the entire system (case 1 versus case 2 in Fig. 1e). For instance, we can validate whether the predator dynamics follow the Lotka–Volterra model regardless of whether the prey dynamics also follow the Lotka–Volterra model (Lotka–Volterra in Fig. 2f) or follow a more complex model (complicated prey dynamics in Fig. 2f). Second, the covariance test is invariant to unknown species that indirectly interact with the species under examination (case 3 in Fig. 1e). For instance, the test for the predator remains valid even when we add many other species that only interact with the prey (species-rich Lotka–Volterra in Fig. 2f), including interactions that ultimately drive chaotic dynamics (chaotic food web in Fig. 2f). Lastly, the covariance criteria can sometimes tolerate unknown abiotic factors (case 4 in Fig. 1e), particularly when the environment acts as an exogenous driver statistically independent of population abundance (with exogenous environment in Fig. 2f).

### Addressing noise and limited data

Ecological time series data are often noisy and limited in length, which can make it difficult to draw reliable inferences from point estimates of covariance values alone. To estimate uncertainty around the covariance measures, we can take a bootstrapping approach. Specifically, we repeatedly draw random samples from the original time series data with replacement and recalculate the gain and loss rate covariances for each resampled dataset. This generates distributions of covariance that capture the inherent variability in the data.

We then compare the distribution of gain rate covariances and the distribution of loss rate covariances. To assess the statistical significance of any observed difference between these distributions, we examine the distribution of their pairwise differences. A significant overlap between the pairwise difference distribution and zero suggests that the model-predicted equality between the covariances is statistically supported. In contrast, a pairwise difference distribution that is clearly shifted away from zero provides strong evidence that the model violates the covariance criteria. To quantify this difference, we calculate Cohen’s  $d$ , a standard measure of effect size between the pairwise difference distribution and zero. A  $z$ -score below a threshold (typically 1.96 for 95% confidence) indicates that the distributions are statistically indistinguishable—the covariances are not different from each other and the model passes the test. This threshold can be adjusted to control the balance between false positives and false negatives as needed. For the planktonic predator–prey example in Fig. 2, the  $z$ -score is 0.04, indicating that the two covariances are probably the same. We have developed the R package ecoModelOracle to streamline this statistical analysis, making it easier for users to implement the approach.

It is important to understand the nature of potential errors that can arise from this use of the covariance criteria to validate or invalidate ecological models. Simulations show that the covariance test is more prone to type I errors (validating an incorrect model) than to type II errors (failing to validate a correct model) (Supplementary Note 3). This is because our test sets a high threshold for validation and therefore minimizes the risk. As a consequence, when a model fails our test, we

can be highly confident in invalidating it. However, when passing the test, the model should be interpreted as a strong candidate for further investigation rather than being the true model.

As a direct application of this statistical approach to the covariance criteria, we revisit the long-standing debate on whether the hare–lynx dynamics in the Canadian boreal zone adhere to the Lotka–Volterra model (Fig. 1e). We do so by applying the covariance criteria to the full dataset from the system. For hares, the calculated  $z$ -score between the distributions of gain and loss covariances is 4.9, a value that strongly suggests unequal covariances and the inadequacy of the Lotka–Volterra model in capturing hare dynamics. In contrast, the  $z$ -score for the gain and loss covariance distributions for lynx is 1.7. This implies that one cannot statistically reject the equality of covariances, supporting the Lotka–Volterra model as a potentially useful approximation for lynx dynamics. These findings mirror case 2 in Fig. 1, aligning with the hypothesis that the Lotka–Volterra model might be valid for lynx but not for hares in this predator–prey system<sup>9</sup>.

### Caveats when applying covariance criteria

While the covariance criteria offer a powerful and broadly applicable tool for rigorously testing ecological models, this is not a panacea. As with any computational method, it is crucial to recognize its limitations and apply it thoughtfully to ensure reliable interpretation.

One key practical consideration arises in cases where not all model parameters cancel algebraically within the covariance equality (equation (3)). As noted earlier, the application of the criteria is most straightforward when parameters do cancel (as in equation (5)), enabling a non-parametric test of the structure of the model using only abundance data. However, more complex models often yield semiparametric tests where some parameters remain within the covariance equality. To deal with this, one can always estimate these values using standard fitting procedures. However, doing so can reintroduce the parameter estimation uncertainties that this method partly aims to avoid. Instead, we recommend evaluating the covariance equality (equation (3)) across a range of biologically realistic values for the necessary parameters. If the equality consistently fails across this plausible parameter space, it provides strong evidence that the fundamental structure of the model is inadequate, irrespective of the precise parameter values. Methods and Supplementary Note 4 illustrate this semiparametric approach by testing two classes of population dynamic models against the global population dynamics database<sup>19</sup>.

Beyond parameter considerations, the interpretation of the covariance criteria hinges on a crucial assumption regarding the primary driver of population fluctuations. The method is designed to test the structure of deterministic gain and loss processes that themselves generate sustained, non-equilibrium dynamics (such as limit cycles or chaos). In these scenarios, the observed variability is a direct reflection of the interplay between the modelled gain and loss rates, making the test powerful and informative. Conversely, the applicability of the method diminishes if the dynamics of the system are governed by a stable equilibrium, where the observed fluctuations arise from stochastic noise perturbing the system away from its equilibrium. While these fluctuations contain information about the stabilizing feedbacks of the system, they do not reflect the specific gain–loss structure that our test interrogates. Applying the criterion in such a noise-dominated scenario, where the deterministic signal is absent, may become uninformative or potentially misleading. Ultimately, the key distinction lies in the origin of the fluctuations: our method is built to analyse deterministic dynamics, not stochastic noise. Rigorous mathematical conditions supporting this point are detailed in Methods, with corresponding simulation analyses in Supplementary Note 5.

These complicating factors, along with other important details concerning the applicability of the covariance criteria, such as the handling of non-stationary data (trends) and the implications of data transformations, are discussed further in Supplementary Note 6.

## Three case studies

In this section, we apply the covariance criteria to tackle three long-standing problems in ecology. We begin by investigating the fundamental nature of predation, using the criteria to rigorously test different functional forms describing the predator–prey interaction. We then study the integration of rapid evolutionary processes into ecological models, leveraging the criteria to potentially pinpoint where evolutionary forces significantly shape species dynamics. Finally, we search for the often-hidden influence of higher-order interactions within ecosystems, harnessing the criteria to uncover complex relationships that extend beyond simple pairwise effects.

### Reverse engineering the nature of predation

Predation is a key force structuring ecological communities<sup>20</sup>. Yet, modelling this interaction remains a challenge. A central debate persists over whether predation is best described by Lotka–Volterra (in which predation is a function of the product of predator and prey abundances), or instead, additionally dependent on the ratio of predator and prey individuals in the system. This debate has persisted for decades<sup>21–25</sup>, partly because traditional methods often struggle to definitively rule out alternative explanations<sup>26</sup>. Here we demonstrate how the covariance criteria can help resolve this long-standing question.

We examine three possible prey dynamics: Lotka–Volterra dynamics with and without self-regulation and ratio-dependent dynamics<sup>27,28</sup>,

$$\frac{dy}{dt} = \begin{cases} \alpha y - \beta yx, & \text{(Lotka–Volterra without self-regulation)} \\ \alpha y - \beta_1 y^2 - \beta_2 yx, & \text{(Lotka–Volterra with self-regulation)} \\ \alpha y - \beta_1 y^2 - \beta_2 \frac{yx}{x+ky} & \text{(Ratio-dependent)} \end{cases} \quad (6)$$

Similarly, we consider three possible predator dynamics:

$$\frac{dx}{dt} = \begin{cases} \delta yx - yx, & \text{(Lotka–Volterra with self-regulation)} \\ \delta yx - \gamma_1 x^2 - \gamma_2 x, & \text{(Lotka–Volterra with self-regulation)} \\ \delta \frac{yx}{x+ky} - yx & \text{(Ratio-dependent)} \end{cases} \quad (7)$$

To rigorously test these models, we analyse a unique long-term dataset with replicated predator–prey dynamics under various conditions<sup>17</sup>. The dataset tracks the relationship between the aquatic invertebrate consumer *Brachionus calyciflorus* and its green algae prey *Monoraphidium minutum*. We find that the prey dynamics align most closely with Lotka–Volterra dynamics with self-regulation (Fig. 3a–c), while the predator dynamics align most closely with Lotka–Volterra dynamics without self-regulation (Fig. 3d–f). For both species, the model with ratio-dependent interactions deviates most from the equality constraint posed by the covariance criteria (see Supplementary Note 7 for additional statistical analysis). Our findings therefore provide compelling evidence that predation in this system is prey-dependent as posed in the Lotka–Volterra model and not a function of the ratio of predators and prey in this system.

To evaluate whether traditional approaches are capable of achieving similar results as the covariance criteria, we applied three distinct methodologies—derivative regression, Bayesian statistics and symbolic regression with deep learning—to the same dataset (Methods). They all fail to capture the inherent dynamics (Supplementary Note 8).

In addition, our analysis sheds light on a long-standing question about where self-regulation emerges in predator–prey systems. We find that self-regulation may play a role in the dynamics of the prey species, but not in the predator species. This observation is consistent with the broader ecological hypothesis that top predators lack strong self-regulating mechanisms<sup>29–32</sup>. One consequence of self-regulation in the prey species, as found here, is an implied role for stochasticity in

shaping the persistent cycles characteristic of predator–prey systems. Without stochasticity, self-regulation within the prey population drives the system towards a stable equilibrium in the Lotka–Volterra model<sup>33</sup>. However, when this self-regulation interacts with environmental stochasticity, the equilibrium is disrupted and transient dynamics can cause indefinite fluctuations<sup>34,35</sup>.

### Dissecting ecological and evolutionary processes

Evolutionary and ecological processes can operate on similar timescales<sup>36,37</sup>. Prey–predator dynamics, in particular, have emerged as a prime example of such rapid evolution<sup>38,39</sup>. However, a major modelling challenge lies in determining where in the ecological system to incorporate evolution: should we focus on prey evolution<sup>40,41</sup>, predator evolution<sup>42</sup> or their simultaneous co-evolution<sup>43,44</sup>. Unfortunately, several models, each incorporating different assumptions about which species evolve, can produce similar observable patterns, including for example the synchrony of predator and prey population cycles<sup>45</sup>. This makes it difficult to pinpoint the specific evolutionary processes operating within the interaction based solely on qualitative observations of the data. Fortunately, the covariance criteria, with its ability to test how well a model captures the key dynamics of each species, offers a promising avenue to pinpoint the specific evolutionary processes at play.

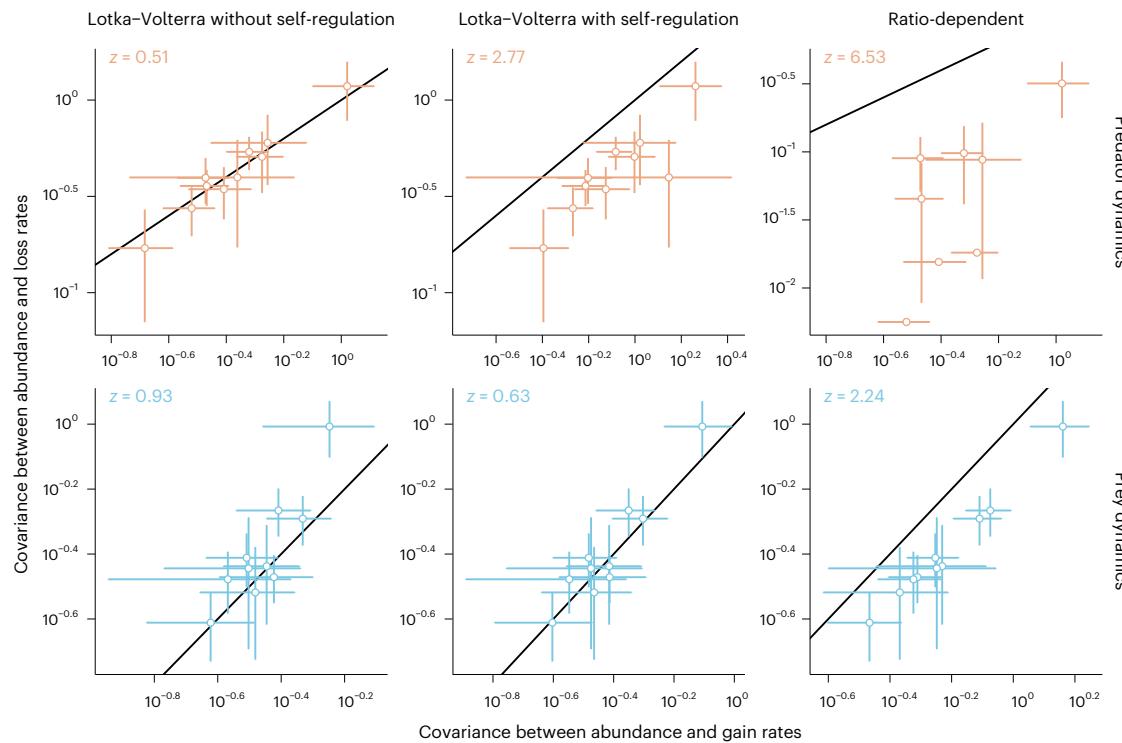
To examine how evolution shapes the dynamics of predator and prey, we must first establish a baseline: how do prey–predator dynamics appear without rapid evolution? Building on our earlier finding (Fig. 3), we propose the Lotka–Volterra model with self-regulation for the prey and without self-regulation for the predator as a candidate. To test the validity of this model, we analysed 18 time series across diverse ecosystems where rapid evolution is not thought to be operating in a major way<sup>17,46–52</sup>, compiled and processed by ref. 53 except for ref. 17 (Supplementary Fig. 20). Applying the covariance criteria to these datasets, we find that the proposed form of the Lotka–Volterra model generally describes both prey and predator dynamics well in the absence of rapid evolution (Fig. 4a,b, statistical analysis in Supplementary Fig. 21).

With a reliable ‘no evolution’ baseline model in hand, we can now ask: how does rapid evolution reshape the covariance structure of the system—in the prey, the predator or both? To address this question, we analysed 13 prey–predator time series where rapid evolution is empirically observed<sup>38,46,52,54–60</sup>. These datasets, compiled and processed by ref. 53, encompass a diverse range of ecosystems, providing an ideal test-bed. The covariance criteria reveal a striking pattern: predator species exhibit significant deviations from the baseline Lotka–Volterra model (with no rapid evolution), suggesting that the Lotka–Volterra model no longer holds (Fig. 4d and Supplementary Fig. 21 in Supplementary Note 9). In contrast, prey species continue to adhere to the predictions of the Lotka–Volterra model (Fig. 4c and Supplementary Fig. 21 in Supplementary Note 9).

These findings suggest that incorporating rapid evolution might require modifications to the predator component of the Lotka–Volterra model, but probably not the prey component. A caveat, though, is that we cannot pinpoint specific evolutionary mechanisms as we have exclusively focused on phenomenological models. It is possible that evolution occurs in the predator’s capture-related traits and/or the prey’s defensive traits, but phenomenologically, only the predator seems to respond to these evolutionary changes in one or both species. Another caveat is that prey species may have different intrinsic gain rates with or without rapid evolution. Owing to the non-parametric nature of the covariance criteria test, we cannot detect these potential differences because they share the same model structure. Despite these limitations, our findings provide guidance for selecting current eco-evolutionary models and catalysing the development of new ones.

### Detecting signals of higher-order interactions

Higher-order interactions, where a third species modifies interactions between a pair, have long fascinated ecologists<sup>61,62</sup>. Yet, detecting their



**Fig. 3 | Reverse engineering the nature of predation.** We apply the covariance criteria to study the functional form of the predation dynamics. We analyse the dataset<sup>17</sup> with ten replicates of an experiment studying the aquatic invertebrate *B. calyculiflorus* (predator) and the green algae *M. minutum* (prey) under varying conditions. Rows represent either prey (blue) or predator (orange) dynamics, while columns compare three models: Lotka–Volterra with self-regulation (centre), Lotka–Volterra without self-regulation (left) and a ratio-dependent model (right). Each panel compares normalized covariances (that is, covariance divided by the mean; equation (3)) between gain/loss rates and abundance (x and y axis,

respectively) and the diagonal line denotes where the two covariances are equal. Data are presented as mean  $\pm$  2 s.d. for each replicate. The value in the upper-left corner of each panel displays the average  $z$ -score of the replicates within that panel. We find that the Lotka–Volterra model with self-regulation (centre) best captures prey dynamics, while the Lotka–Volterra model without self-regulation (left) best describes predator dynamics (Supplementary Figs. 15 and 16 provide statistical details). These findings suggest that a prey-dependent functional form, as used in the Lotka–Volterra model, is more appropriate to describe predation in this system compared with a ratio-dependent model.

existence remains challenging. Experimental manipulations, while ideal, are often logistically difficult<sup>63,64</sup>. A common alternative is to infer higher-order interactions through model fitting<sup>65–67</sup>. However, since higher-order interactions introduce more parameters, models can overfit the data, giving the illusion of higher-order interactions where none exists<sup>68,69</sup>. While regularization methods and information criteria can mitigate this issue<sup>70,71</sup>, biases may still persist.

In contrast, the covariance criteria offer a compelling alternative for detecting potential higher-order interactions, as they are inherently less susceptible to overfitting. Specifically, higher-order interactions, when encoded in a model, change the predicted covariance structure. If that model was applied to a dataset with no true higher-order interactions, a mismatch between the predictions of the model and the observed data would emerge. To demonstrate this, we analyse a high-quality, long-term dataset of a rocky intertidal community in Goat Island Bay, New Zealand<sup>72</sup>. This dataset tracks the monthly percentage cover of barnacles, mussels and algae for over 20 years. Others<sup>72</sup> proposed a model without higher-order interactions for mussel dynamics: where  $M$  is the cover of mussels,  $B$  is the cover of barnacles,

$$\frac{dM}{dt} = r(A + B)M - zF(t)M \quad (8)$$

Mussel  
Crustose algae   Barnacle   Seasonality effect =  $1 + \alpha(T_{\max} - T_{\text{mean}}) \cos\left(\frac{2\pi(t-32)}{365}\right)$

$A$  is the cover of crustose algae,  $r$  is the rate at which area covered by those two species is colonized by mussels,  $z$  is the constant death rate of mussels and  $F(t)$  represents the effects of seasonality, which is a complex function of abiotic factors. The model is formulated to

indicate that mussels require cover by barnacles and/or algae to colonize the marine intertidal.

Despite the complexity of this model, it has a simple covariance structure

$$\frac{\text{Cov}(r(A + B)M, M)}{\langle r(A + B)M \rangle \langle M \rangle} = \frac{\text{Cov}(zF(t)M, M)}{\langle zF(t)M \rangle \langle M \rangle} \quad (9)$$

$$\Rightarrow \frac{\text{Cov}((A + B)M, M)}{\langle (A + B)M \rangle \langle M \rangle} = \frac{\text{Cov}(M, M)}{\langle M \rangle \langle M \rangle} \quad (10)$$

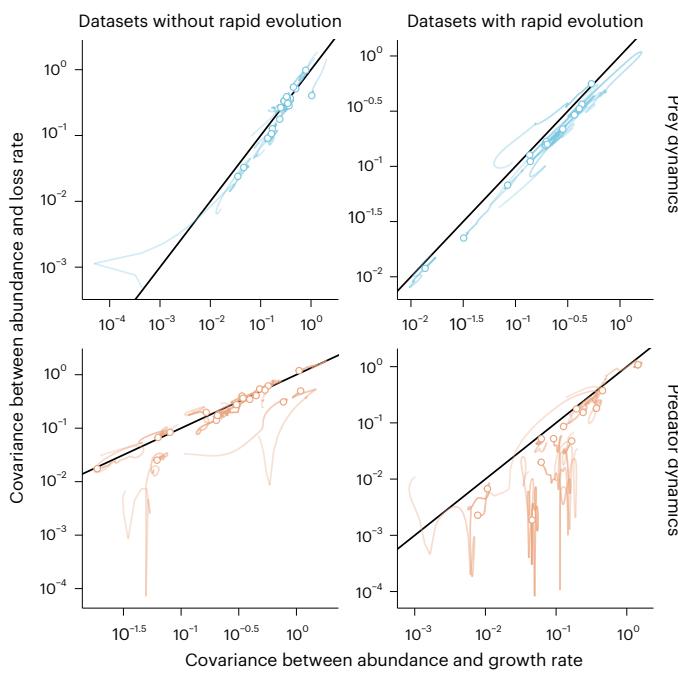
We can cancel the mussel colonization ( $r$ ) and death rate ( $z$ ) because they are constant and can cancel the effects of seasonality  $F(t)$  because  $F(t)$  is independent of the fluctuations of mussels  $M$  ( $P = 0.72$  with non-linear correlation test<sup>73</sup>).

Additionally, we considered two further models. One model assumes that mussel growth depends only on a higher-order interaction—the interactive effect of algae and barnacles on mussel colonization:

$$\frac{dM}{dt} = r \frac{ABM}{dt} - zF(t)M, \quad (11)$$

Higher order only

and the other model combines the pairwise and higher-order interactions:



**Fig. 4 | Dissecting ecological and evolutionary processes.** We apply the covariance criteria to study how rapid evolution affects population dynamics in predator–prey systems. Inspired by previous analyses, we evaluate the Lotka–Volterra model with self-regulating prey. Rows represent either prey (blue) or predator (orange) dynamics. Columns differentiate between datasets with (right panels; 13 datasets; 18 datasets) evidence of rapid evolution. Each panel compares normalized covariances (that is, covariance divided by the mean; equation (3)) between gain/loss rates and abundance (x and y axis, respectively) and the diagonal line denotes where the two covariances are equal. Each line represents the results of the covariance criteria test applied to a specific dataset, using a different time window within that dataset. The transparency of the line indicates the size of the time window used: less transparent lines signify longer length of the time series segments starting at the origin of the time series, while a dot represents the analysis using the full time range of the dataset. We find that, without rapid evolution (left panels), the Lotka–Volterra model effectively describes both prey and predator dynamics across ecosystems. In contrast, with rapid evolution (right panels), the Lotka–Volterra model remains suitable for prey dynamics but not for predator dynamics. Supplementary Fig. 21 gives statistical details. These results guide how to incorporate rapid evolution in modelling prey–predator dynamics.

$$\frac{dM}{dt} = r (A + B + AB)M - zF(t)M \quad (12)$$

↑  
Pairwise + Higher-order

We then test the covariance structure from all three models above (equations (8), (11) and (12)) against the empirical data (Fig. 5). We find that the higher-order interaction-only model (equation (11)) fails the test entirely ( $z = 6.17$ , gain rate covariance with density is quite different from the loss rate covariance). In contrast, the pairwise only model (equation (8)) shows similar covariances for the gain and loss rates ( $z = 1.75$ ). The pairwise + higher-order interaction model (equation (12)) almost perfectly explains the data ( $z = 0.09$ ), suggesting the presence of higher-order interactions within this system (Supplementary Note 10 provides further statistical analysis). Such higher-order interactions could involve the barnacles and algae covered areas synergistically facilitating mussel colonization.

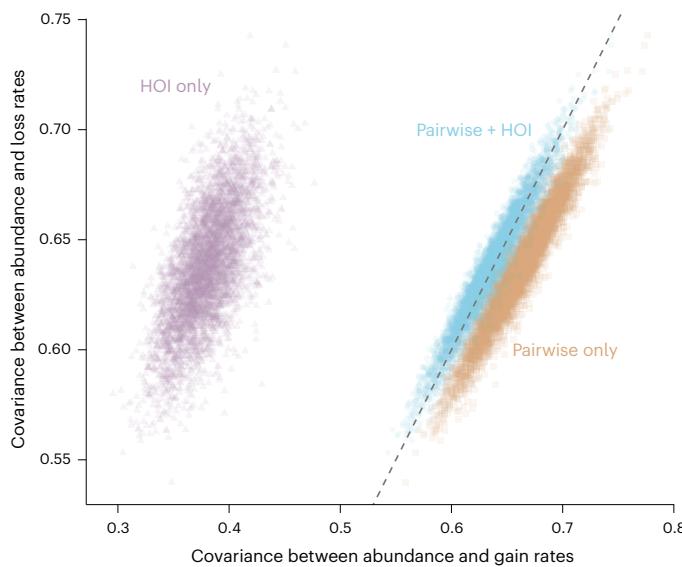
## Discussion

We introduce the covariance criteria as a powerful, assumption-light framework for validating ecological models against time series data. The key insight is that every dynamical model imposes unique constraints on the permissible covariance structures relating population abundances, gain rates and loss rates. If the empirical data satisfy these constraints, we can be confident the model provides a useful approximation capturing core aspects of the dynamics of the system. Conversely, violations of the covariance criteria provide quantitative evidence that the model is fundamentally inadequate, at least for the particular species and conditions examined.

Theoretically, the covariance criteria exhibit remarkable generality, applying across ecological dynamics ranging from simple equilibria to complex non-equilibrium systems with non-Markovian delays and external stochastic forcing. Computationally, the criteria are efficient to evaluate and often operate non-parametrically (or semiparametrically), eliminating the need to specify all model parameters from data. Perhaps most crucially from an empirical standpoint, the covariance criteria can be readily applied to the limited and noisy time series data common in ecological studies. As demonstrated through our three case studies, this approach consistently supports ecological models aligning with prevailing ecological understanding, while decisively rejecting those failing to capture underlying dynamics. In an era of rapidly accumulating high-quality ecological data, this approach subjects theorists' ideas to rigorous scrutiny and facilitates a better dialogue between ecological theory and empirical reality.

Theoretical ecologists have often been criticized for validating models with a low bar for consistency with data<sup>74–76</sup>. The low validation bar allows a multiplicity of models to appear acceptable, even when their predicted mechanisms are vastly different, leading to insufficient confidence in any particular model. However, this raises a question: if we set a more rigorous quantitative bar, would all ecological models fail? This concern may explain the limited attention the covariance criteria have received beyond its originators<sup>77–79</sup>. After all, as R. May wrote, “the models of biological communities tend rather to be of a very general, strategic kind”<sup>80</sup>. We initially expected most ecological models to struggle to meet the strict covariance criteria. Much to our surprise, however, we found that the classic Lotka–Volterra model withstood the test across a wide range of consumer–resource systems. This stands in direct contrast to the common perception, echoed in many introductory ecology texts<sup>26,39</sup>, that the Lotka–Volterra model is overly simplistic and misses crucial biological details. Our findings could help explain the recent success of the Lotka–Volterra model in predicting some ecological patterns<sup>81–83</sup>.

This surprising robustness of the Lotka–Volterra model under our test of the covariance criteria, especially given its perceived simplicity, raises the question of why traditional validation methods have often failed to support it in the same systems. Understanding this discrepancy requires considering how different approaches handle the complexities introduced by noise in real-world ecological data. While system identification theory confirms that Lotka–Volterra models are uniquely identifiable under idealized, noise-free conditions<sup>84,85</sup>, ecological systems are rarely noise-free. Stochastic perturbations often introduce more than simple measurement error, potentially causing subtle but significant effects such as transient phase shifts in population cycles. Traditional validation methods, which often rely heavily on capturing the precise timing and amplitude of oscillations or assume simpler noise structures, can be highly sensitive to these disruptions. Consequently, they might incorrectly reject an underlying Lotka–Volterra structure due to noise-induced deviations from deterministic expectations. In contrast, the covariance criteria approach proves more resilient because it focuses on arguably the simplest statistics—specifically means and covariances—averaged over time. These properties are inherently less sensitive to the exact phasing or amplitude of individual cycles, allowing our method to capture the



**Fig. 5 | Detecting signals of higher-order interactions.** We apply the covariance criteria to study the presence of higher-order interactions (HOI) in a rocky intertidal community<sup>72</sup>. Specifically, whether crustose algae and barnacles interact with mussels exclusively through pairwise interactions or whether HOI is present. Three models are evaluated: pairwise interactions only (orange), higher-order interaction-only (purple) and a combination of both (blue). The x axis represents the covariance between abundance and gain rates, while the y axis represents the covariance between abundance and loss rates. Points are derived from 1,000 bootstrapping replicates. The higher-order only model (purple) shows a significant mismatch in covariance values, indicating its inadequacy. The pairwise interaction model (orange) aligns more closely but still deviates statistically from the observed loss covariance. In contrast, the model incorporating both pairwise and HOI (blue) accurately captures the loss covariance. Supplementary Fig. 23 shows further statistical analysis. These findings strongly suggest that both pairwise and HOI between crustose algae, barnacles and mussels play a significant role in influencing mussel dynamics within this community.

fundamental signature of the underlying ecological dynamics even amid the complexities of realistic noise and perturbations.

The covariance criteria represent a fundamentally different approach to model validation than machine learning methods such as symbolic regression<sup>86–88</sup>. Those machine learning techniques aim to directly distil mathematical models from patterns in empirical data, with minimal a priori knowledge of the system. In contrast, the covariance criteria approach retains theory-based model building at its core, as it begins with a theorist-proposed dynamical model inspired by natural history. Importantly, covariance criteria and machine learning approaches can work together. For example, they could be combined into a synergistic modelling pipeline where machine learning suggests new model structures, theorists use their expertise to explain the mechanisms and the covariance criteria rigorously test the resulting models against new data.

The covariance criteria could be useful for evaluating more than just population dynamics over time, as we focused on here. For example, the criteria could be used to test models for how population abundances vary over space, analogous to common ecological approaches of substituting space for time when aiming to understand long-term dynamics<sup>89,90</sup>. However, doing so would require making additional assumptions about how model parameters vary across locations. Future work could also adapt the criteria to model the dynamics of other types of empirical data, such as temporal changes in trait values or single nucleotide polymorphisms.

Despite their advantages, the covariance criteria have limitations. The approach is most effective when a species has only few direct

interactions with other species. This is because the criteria partition the model into gain and loss components without dissecting their underlying process in any detail. When species interact directly with many rather than a few species, the gain and loss terms can become highly complex, parameters often can be cancelled out or systematically varied. While the covariance approach is still theoretically applicable, it requires careful consideration of parameter ranges and may not be feasible.

As ecology grapples with increasingly complex challenges, from climate change to biodiversity loss, the need for reliable models has never been greater. The covariance criteria approach offers a path towards greater confidence in our ecological understanding by rigorously testing models across a wide range of problems. By providing a more rigorous foundation for model validation, we hope this method can contribute to more accurate predictions of ecosystem responses to environmental perturbations and more effective conservation strategies.

## Methods

### Partitioning into gain and loss rates

Our approach partitions any given dynamical equation in the most direct and mechanistically plausible way: the total gain rate ( $R_+$ ) is the sum of all positive terms and the total loss rate ( $R_-$ ) is the sum of the absolute values of all negative terms. As an example, we partition the logistic equation,  $\frac{dN}{dt} = rN - \alpha N^2$ , into a gain rate  $R_+ = rN$  (representing linear, density-independent gain) and a loss rate  $R_- = \alpha N^2$  (representing quadratic, density-dependent loss).

This partitioning is not arbitrary; it is constrained by fundamental requirements of the underlying theory. To illustrate, consider an alternative but ultimately invalid partitioning of the same logistic equation, where the gain rate is  $R'_+ = (r+k)N - \alpha N^2$  (representing nonlinear, density-dependent gain) and the loss rate is  $R'_- = kN$  (representing linear, density-independent loss) for some positive constant  $k$ . This formulation is invalid within our framework for two primary reasons as described next.

First, gain or loss rates cannot be negative. The proposed gain term,  $R'_+ = (r+k)N - \alpha N^2$ , is a parabola that becomes negative for large  $N$ , which is biologically and mathematically nonsensical. One might attempt to resolve this by defining the gain rate with a floor at zero, that is,  $R''_+ = \max(0, (r+k)N - \alpha N^2)$ . However, this fundamentally alters the dynamics of the system. The resulting model,

$$\frac{dN}{dt} = \max(0, (r+k)N - \alpha N^2) - kN \quad (13)$$

is now a new and distinct model, not a repartitioning of the classic logistic equation. This is immediately evident as its ecological equilibrium (where gain equals loss) is no longer at  $N' = r/\alpha$ .

Second, the covariance criteria are derived assuming that the stochastic noise affecting the overall gain process is statistically independent of the noise affecting the loss process. By defining the gain rate as containing a term ' $+kN$ ' and the loss rate as being equal to ' $kN$ ', we imply that the a set of events responsible for the entire loss process is also a component of the gain process. This introduces a correlation between a component of the gain fluctuations and the entirety of the loss fluctuations, violating a key assumption of the theoretical framework.

More importantly, our method is designed to test a specific proposed phenomenological model. The standard logistic model is the hypothesis that gain is linear and loss is quadratic. The alternative scenario—for instance, one with a saturating gain rate and a linear loss rate—is a different phenomenological model entirely, better captured by an equation such as  $\frac{dN}{dt} = \frac{rN}{1+cN} - dN$ . Each distinct partitioning of gain and loss terms constitutes a unique, testable hypothesis about the underlying structure of the system.

### Illustrative simulation-based validation

The covariance constraint (equation (3)) is a mathematically derived equality. To intuitively illustrate its robustness to unmodelled ecological complexity, we performed simulations in Fig. 2f. In these simulations, we fix the assumed structure for the predator's abundance ( $x$ ). We then vary the underlying dynamics of the prey ( $s$ ) or introduce other complexities into the system, demonstrating that the equality between the normalized covariances holds regardless of these external dynamics.

The first dynamic simulated corresponds to the classic Lotka–Volterra predator–prey model:

$$\frac{ds}{dt} = \alpha s - \beta sx \quad (14)$$

$$\frac{dx}{dt} = \delta sx - \gamma x \quad (15)$$

where  $\alpha$  is the intrinsic growth rate of the prey,  $\beta$  is the predation rate (effect of predator on prey growth),  $\delta$  is the conversion efficiency (effect of prey on predator growth) or rate constant for predator growth and  $\gamma$  is the predator mortality rate.

To add more complexity, the second simulated dynamic includes a type-2 functional response for the prey,

$$\frac{ds}{dt} = \alpha \left( s - \frac{s}{K} \right) - \frac{\beta sx}{1 + s} \quad (16)$$

$$\frac{dx}{dt} = \delta sx - \gamma x \quad (17)$$

where  $\alpha$  is the intrinsic growth rate of the prey,  $K$  is the carrying capacity of the environment for the prey,  $\beta$  is the maximum predation rate or attack rate.

To complicate the dynamical behaviour, the third dynamic is a chaotic food web model involving a resource ( $s_1$ ), an intermediate prey ( $s_2$ ) and the predator ( $x$ ) feeding on the intermediate prey<sup>91</sup>:

$$\frac{ds_1}{dt} = rs_1 \left( 1 - \frac{s_1}{k} \right) - \frac{a_1 s_1 x}{1 + b_1 s_1} \quad (18)$$

$$\frac{ds_2}{dt} = -c_1 s_2 + c_2 s_1 s_2 - \frac{a_2 s_2 x}{1 + b_2 s_2} \quad (19)$$

$$\frac{dx}{dt} = \delta s_2 x - \gamma x \quad (20)$$

where  $a_1$  is the attack rate of the intermediate prey on the resource,  $b_1$  is the handling time of the intermediate prey for the resource,  $c_1$  is the mortality rate of the intermediate prey,  $c_2$  is the conversion efficiency of the resource into intermediate prey,  $a_2$  is the attack rate of the predator on the intermediate prey,  $b_2$  is the handling time of the predator for the intermediate prey.

To introduce environmental variation, we considered a scenario with an exogenous, time-varying environment affecting the predator's mortality rate:

$$\frac{ds}{dt} = \alpha s - \beta sx \quad (21)$$

$$\frac{dx}{dt} = \delta sx - \cos(t) \gamma x \quad (22)$$

where  $\gamma$  is the baseline predator mortality rate and  $\cos(t)$  models the seasonal variation of mortality.

Finally, to explore the effect of network complexity, we simulated a species-rich system where the prey dynamics ( $s$ ) are coupled to a hundred other species ( $s_j$ ):

$$\frac{ds}{dt} = rs + \sum_{j=1}^{100} a_j s_j \quad (23)$$

$$\frac{dx}{dt} = \delta sx - \gamma x \quad (24)$$

where  $a_j$  is the interaction strength between the focal prey species ( $s$ ) and the  $j$ th other species ( $s_j$ ).

As demonstrated visually in Fig. 2f, the calculated points for each model fall along the 1:1 line in Fig. 2f. This visually reinforces that the covariance criterion for the predator  $x$  (with assumed gain rate  $R_+$  and loss rate  $R_-$ ) holds across these diverse ecological scenarios, highlighting its robustness to many forms of unmodelled dynamics external to the specific gain–loss structure being tested.

### Application to semiparametric models

As discussed in the main section on 'Caveats when applying covariance criteria', the covariance criteria test may become semiparametric when model parameters do not cancel algebraically. In such cases, we advocate evaluating the covariance equality (equation (3)) across a plausible range of the non-cancellable parameter(s). To demonstrate this concretely, we performed a meta-analysis using the high-quality subset of a large dataset of single-species dynamics. The global population dynamics database is currently the largest compilation of time series for single species<sup>19</sup>. As some data are either too short or too noisy for robust time series analysis, we restricted our analysis to the subset selected by ref. 92, which contains 172 time series spanning 138 different taxa and 57 sampling locations.

We tested two common single-species growth models where parameters do not fully cancel. The first one is the theta logistic growth model<sup>93,94</sup>

$$\frac{dN}{dt} = rN \left( 1 - \left( \frac{N}{K} \right)^\theta \right) \quad (25)$$

where  $\theta > 1$  is the non-cancellable parameter. The second is the sublinear growth model<sup>95,96</sup>:

$$\frac{dN}{dt} = rN^k - bN \quad (26)$$

where the parameter  $0 < k < 1$  is non-cancellable.

For each model, we systematically varied the non-cancellable parameter ( $\theta$  or  $k$ ) across its biologically relevant range. We found that these models consistently failed the covariance test across the parameter ranges for almost all datasets, providing strong evidence for their structural inadequacy in describing these populations. Detailed results and specific parameter ranges explored are presented in Supplementary Note 4.

### Nature of fluctuations in covariance criteria

The covariance criteria assume that fluctuations are primarily driven by the deterministic gain ( $R_+$ ) and loss ( $R_-$ ) processes. To examine how additive noise might affect this, we formalize a general scenario where the rates include noise:  $R'_+ = R_+ + \epsilon_+$  and  $R'_- = R_- + \epsilon_-$ , where the terms  $\epsilon_+$  and  $\epsilon_-$  are stochastic noise impacting the gain and loss processes. We assume these are independent noise processes, each with zero mean. Crucially, this general formulation allows the noise affecting gain ( $\epsilon_+$ ) to differ from that affecting loss ( $\epsilon_-$ ) and permits both to be density-dependent (functions of abundance  $x$ ) or density-independent.

The ratio of covariances in the presence of noise then becomes: Next we discuss when the new constraint (equation (27)) reduces to the original constraint (equation (3)):

$$\begin{aligned} & \left( \frac{\text{Cov}(R_+, \epsilon_+, x)}{(R_+ + \epsilon_+, x)} \right) / \left( \frac{\text{Cov}(R_-, \epsilon_-, x)}{(R_- + \epsilon_-, x)} \right) \\ &= \left( \frac{\text{Cov}(R_+, x) + \text{Cov}(\epsilon_+, x)}{(R_+ + \epsilon_+)} \right) / \left( \frac{\text{Cov}(R_-, x) + \text{Cov}(\epsilon_-, x)}{(R_- + \epsilon_-)} \right) \quad \text{Mean of noise is zero} \\ &= \left( \frac{\text{Cov}(R_+, x) + \text{Cov}(\epsilon_+, x)}{(R_+)} \right) / \left( \frac{\text{Cov}(R_-, x) + \text{Cov}(\epsilon_-, x)}{(R_-)} \right) \quad \text{Linearity of covariance and mean} \end{aligned} \quad (27)$$

**Density-independent noise.** If the noise  $\epsilon$  is density-independent, then  $\text{Cov}(R_+, x) = \text{Cov}(R_-, x) = 0$  because the noise is uncorrelated with the population size  $x$ . In this case, the covariance ratio simplifies back to the original noise-free condition (equation (3)) and the covariance criterion remains valid.

**Density-dependent noise.** When the noise  $\epsilon$  is density-dependent,  $\text{Cov}(R_+, x)$  and  $\text{Cov}(R_-, x)$  are not zero and can significantly affect the covariance ratio. The impact depends on the relative magnitude of these covariance terms.

- When gain/loss processes dominate noise, meaning that  $\text{Cov}(R_+, x) \gg \text{Cov}(\epsilon_+, x)$  and  $\text{Cov}(R_-, x) \gg \text{Cov}(\epsilon_-, x)$ , then the additional covariance terms due to noise are negligible and the covariance criterion still holds approximately (because  $\text{Cov}(R_+, x) + \text{Cov}(\epsilon_+, x) \approx \text{Cov}(R_+, x)$ ).
- In contrast, if the density-dependent noise contributes significantly to the fluctuations, the original covariance criterion no longer applies. In this case, the noise must be explicitly accounted for in the analysis. However, the form of density-dependent noise is often unknown or difficult to characterize in ecology. Without specific knowledge of how noise depends on population size, it is challenging to adjust the covariance criterion to account for it. This limitation makes it difficult to apply the criterion accurately when density-dependent noise plays a significant role.

All the arguments above are mathematically rigorous. To see how these principles apply in practice, we performed a detailed simulation study (Supplementary Note 5). The results confirm that the validity of our method depends on the balance between the deterministic ‘signal’ and the stochastic ‘noise’. Specifically, our simulations show that the covariance criteria are extremely robust to additive (density-independent) noise. For density-dependent noise (both multiplicative and demographic), the test performs reliably at low to moderate intensities. Its accuracy only diminishes when the noise becomes strong enough to dominate the dynamics of the system and obscure the deterministic signal.

### Implementation of comparative methods

To benchmark the covariance criteria against standard model validation techniques, particularly for the predation functional response case study (Fig. 3), we applied three common approaches to this dataset<sup>17</sup>. The goal was to assess if these methods could distinguish between the competing functional response models as effectively as the covariance criteria.

**Regression on inferred derivatives.** This method involves linear regression on the inferred derivatives of the population abundances, leveraging the linear relationships derived from transforming the generalized Lotka–Volterra equations. Specifically, the Lotka–Volterra equations for species ( $x_i$ ) can be written as

$$\frac{d \log x_i}{dt} = r_i + \sum_j a_{ij} x_j \quad (28)$$

where  $r_i$  is the intrinsic growth rate and  $a_{ij}$  are interaction strength

between species  $x_i$  and  $x_j$ . This equation shows the linear relationship between the rate of change of the logarithm of abundances and the abundances of the interacting species. This relationship makes a linear regression possible. We then easily set up linear regression models for each species. The key challenge is inferring the derivatives  $\frac{d}{dt} \log x_i$  numerically from noisy data. A naive finite difference approximation

$$\frac{d \log x_i}{dt} \approx \frac{\log(x_i(t_{k+1})) - \log(x_i(t_k))}{t_{k+1} - t_k} \quad (29)$$

amplifies noise, leading to unreliable estimates. To mitigate this, we used the R package `gauseR`, which provides optimized functions for derivative estimation using smoothing techniques.

**Bayesian nonlinear ODEs modelling.** While regression fitting approaches are useful, they might still be biased as they lack the direct modelling of ordinary differential equations (ODEs). Therefore, we also implemented a Bayesian framework. Instead of directly fitting the parameters, we used an iterative process: we start with an initial guess for the parameters, run a simulation of the Lotka–Volterra model and then update the parameters on the basis of the comparison between the simulation output and the observed data. This process is repeated until the model converges to a good fit. Specifically, we used Markov chain Monte Carlo methods to sample from the posterior distribution of the parameters, allowing us to quantify uncertainty in the parameter estimates. We also incorporated uncertainties in the initial abundance measurements, maximizing the use of available information. In our Bayesian model, we specified the Lotka–Volterra dynamics within the Stan language using the `brms` package in R, which facilitates fitting Bayesian models with complex hierarchical structures and custom likelihoods. We then used the `ode rk45` solver to numerically integrate the ODEs over time.

**Symbolic regression with deep learning.** Symbolic regression aims to infer the underlying equations governing a system directly from data, without assuming a specific model form. This approach has gained attention in ecology<sup>36,37</sup>. However, traditional symbolic regression methods are often sensitive to noise, which is prevalent in ecological time series. To address this, we used a state-of-the-art method of symbolic regression via deep learning<sup>37</sup>. This method uses a transformer neural network architecture, similar to those used in large language models, to learn the mathematical relationships between variables. The trained neural network, titled `odeformer`, has a total parameter count of 86 million and was trained on a dataset of 50 million samples of diverse ODEs.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

Empirical data of aquatic invertebrate and the green algae are available from [www.nature.com/articles/s41586-019-1857-0](https://www.nature.com/articles/s41586-019-1857-0). Empirical dataset of consumer–resource dynamics is available from <https://onlinelibrary.wiley.com/doi/full/10.1111/ele.12291>.

### Code availability

The R package `ecoModelOracle` to run the analysis is available via GitHub at <https://github.com/clsong/ecoModelOracle>.

### References

1. Kingsland, S. in *Foundations of Ecology: Classic Papers with Commentaries* (eds Real, L. & Brown, J.) 1–13 (Univ. Chicago Press, 1991).
2. Grainger, T. N. et al. An empiricist’s guide to using ecological theory. *Am. Nat.* **199**, 1–20 (2022).

3. Novak, M. & Stouffer, D. B. Geometric complexity and the information-theoretic comparison of functional-response models. *Front. Ecol. Evol.* **9**, 740362 (2021).
4. Leigh, E. in *Some Mathematical Problems in Biology* (ed. Gerstenhaber, M.) 1–61 (American Mathematical Society, Providence, 1968).
5. Volterra, V. Fluctuations in the abundance of a species considered mathematically. *Nature* **118**, 558–560 (1926).
6. Rosenzweig, M. L. & MacArthur, R. H. Graphical representation and stability conditions of predator–prey interactions. *Am. Nat.* **97**, 209–223 (1963).
7. Gilpin, M. E. Do hares eat lynx? *Am. Nat.* **107**, 727–730 (1973).
8. Cortez, M. H. & Weitz, J. S. Coevolution can reverse predator–prey cycles. *Proc. Natl Acad. Sci. USA* **111**, 7486–7491 (2014).
9. Stenseth, N. C., Falck, W., Bjørnstad, O. N. & Krebs, C. J. Population regulation in snowshoe hare and canadian lynx: asymmetric food web configurations between hare and lynx. *Proc. Natl Acad. Sci. USA* **94**, 5147–5152 (1997).
10. Blasius, B., Huppert, A. & Stone, L. Complex dynamics and phase synchronization in spatially extended ecological systems. *Nature* **399**, 354–359 (1999).
11. Hone, J., Krebs, C. J. & O'Donoghue, M. Is the relationship between predator and prey abundances related to climate for lynx and snowshoe hares? *Wildl. Res.* **38**, 419–425 (2011).
12. Yan, C., Stenseth, N. C., Krebs, C. J. & Zhang, Z. Linking climate change to population cycles of hares and lynx. *Glob. Change Biol.* **19**, 3263–3271 (2013).
13. King, A. A. & Schaffer, W. M. The geometry of a population cycle: a mechanistic model of snowshoe hare demography. *Ecology* **82**, 814–830 (2001).
14. Hilfinger, A., Norman, T. M., Vinnicombe, G. & Paulsson, J. Constraints on fluctuations in sparsely characterized biological systems. *Phys. Rev. Lett.* **116**, 058101 (2016).
15. Little, J. D. A proof for the queuing formula:  $L = \lambda w$ . *Oper. Res.* **9**, 383–387 (1961).
16. Little, J. D. Little's law as viewed on its 50th anniversary. *Oper. Res.* **59**, 536–549 (2011).
17. Blasius, B., Rudolf, L., Weithoff, G., Gaedke, U. & Fussmann, G. F. Long-term cyclic persistence in an experimental predator–prey system. *Nature* **577**, 226–230 (2020).
18. Hilfinger, A., Norman, T. M. & Paulsson, J. Exploiting natural fluctuations to identify kinetic mechanisms in sparsely characterized systems. *Cell Syst.* **2**, 251–259 (2016).
19. Prendergast, J. et al. The Global Population Dynamics Database (knb, 2010); <https://knb.ecoinformatics.org/view/doi:10.5063/F1BZ63Z8>
20. Pringle, R. M. et al. Predator-induced collapse of niche structure and species coexistence. *Nature* **570**, 58–64 (2019).
21. Abrams, P. A. & Ginzburg, L. R. The nature of predation: prey dependent, ratio dependent or neither? *Trends Ecol. Evol.* **15**, 337–341 (2000).
22. Arditi, R. & Ginzburg, L. R. *How Species Interact: Altering the Standard View on Trophic Ecology* (Oxford Univ. Press, 2012).
23. Abrams, P. A. Why ratio dependence is (still) a bad model of predation. *Biol. Rev.* **90**, 794–814 (2015).
24. Tyutyunov, Y. V. & Titova, L. From Lotka–Volterra to Arditi–Ginzburg: 90 years of evolving trophic functions. *Biol. Bull. Rev.* **10**, 167–185 (2020).
25. Ginzburg, L. & Damuth, J. The issue isn't which model of consumer interference is right, but which one is least wrong. *Front. Ecol. Evol.* **10**, 860542 (2022).
26. Morin, P. J. *Community Ecology* (John Wiley & Sons, 2009).
27. Akcakaya, H. R., Arditi, R. & Ginzburg, L. R. Ratio-dependent predation: an abstraction that works. *Ecology* **76**, 995–1004 (1995).
28. Arditi, R. & Ginzburg, L. R. Coupling in predator–prey dynamics: ratio-dependence. *J. Theor. Biol.* **139**, 311–326 (1989).
29. Pimm, S. & Lawton, J. Number of trophic levels in ecological communities. *Nature* **268**, 329–331 (1977).
30. Tilman, D. *Resource Competition and Community Structure* (Princeton Univ. Press, 1982).
31. Chesson, P. in *Ecological Systems: Selected Entries from the Encyclopedia of Sustainability Science and Technology* (ed. Leemans, R.) 223–256 (Springer, 2013).
32. Song, C. & Saavedra, S. Bridging parametric and nonparametric measures of species interactions unveils new insights of non-equilibrium dynamics. *Oikos* **130**, 1027–1034 (2021).
33. Murdoch, W. W., Briggs, C. J. & Nisbet, R. M. *Consumer–Resource Dynamics* MPB-36 (Princeton Univ. Press, 2013).
34. Gurney, W. & Nisbet, R. Single-species population fluctuations in patchy environments. *Am. Nat.* **112**, 1075–1090 (1978).
35. Nisbet, R. M. & Gurney, W. *Modelling Fluctuating Populations* (Blackburn Press, 2003).
36. Schoener, T. W. The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. *Science* **331**, 426–429 (2011).
37. Loreau, M., Jarne, P. & Martiny, J. B. Opportunities to advance the synthesis of ecology and evolution. *Ecol. Lett.* **26**, S11–S15 (2023).
38. Yoshida, T., Jones, L. E., Ellner, S. P., Fussmann, G. F. & Hairston, N. G. Rapid evolution drives ecological dynamics in a predator–prey system. *Nature* **424**, 303–306 (2003).
39. Mittelbach, G. G. & McGill, B. J. *Community Ecology* (Oxford Univ. Press, 2019).
40. Marrow, P., Law, R. & Cannings, C. The coevolution of predator–prey interactions: ESS and Red Queen dynamics. *Proc. R. Soc. Lond. B* **250**, 133–141 (1992).
41. Abrams, P. A. Prey evolution as a cause of predator–prey cycles. *Evolution* **51**, 1740–1748 (1997).
42. Abrams, P. A. Adaptive foraging by predators as a cause of predator–prey cycles. *Evol. Ecol.* **6**, 56–72 (1992).
43. Dieckmann, U. & Law, R. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.* **34**, 579–612 (1996).
44. Gavrillets, S. Coevolutionary chase in exploiter–victim systems with polygenic characters. *J. Theor. Biol.* **186**, 527–534 (1997).
45. Abrams, P. A. The evolution of predator–prey interactions: theory and evidence. *Annu. Rev. Ecol. Syst.* **31**, 79–105 (2000).
46. Utida, S. Population fluctuation, an experimental and theoretical approach. *Cold Spring Harbor Symp. Quant. Biol.* **22**, 139–151 (1957).
47. Huffaker, C. et al. Experimental studies on predation: dispersion factors and predator–prey oscillations. *Hilgardia* **27**, 343–383 (1958).
48. Barnet, Y. M., Daft, M. & Stewart, W. Cyanobacteria–cyanophage interactions in continuous culture. *J. Appl. Bacteriol.* **51**, 541–552 (1981).
49. Dulos, E. & Marchand, A. Oscillations of the population densities of the bacterial prey–predator couple *Escherichia coli*–*Bdellovibrio bacteriovorus*: experimental study and theoretical model. *Ann. Inst. Pasteur/Microbiol.* **135**, 271–293 (1984).
50. Luckinbill, L. S. Coexistence in laboratory populations of *Paramecium aurelia* and its predator *Didinium nasutum*. *Ecology* **54**, 1320–1327 (1973).
51. Luckinbill, L. S. The effects of space and enrichment on a predator–prey system. *Ecology* **55**, 1142–1147 (1974).
52. Veilleux, B. *The Analysis of a Predatory Interaction between Didinium and Paramecium*. MSc thesis, Univ. Alberta (1976).
53. Hiltunen, T., Hairston Jr, N. G., Hooker, G., Jones, L. E. & Ellner, S. P. A newly discovered role of evolution in previously published consumer–resource dynamics. *Ecol. Lett.* **17**, 915–923 (2014).

54. Tsuchiya, H., Drake, J., Jost, J. & Fredrickson, A. Predator-prey interactions of *Dictyostelium discoideum* and *Escherichia coli* in continuous culture. *J. Bacteriol.* **110**, 1147–1153 (1972).

55. Jost, J., Drake, J., Fredrickson, A. & Tsuchiya, H. Interactions of *Tetrahymena pyriformis*, *Escherichia coli*, *Azotobacter vinelandii*, and glucose in a minimal medium. *J. Bacteriol.* **113**, 834–840 (1973).

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

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

58. Boraas, M. E. in *Evolution and Ecology of Zooplankton Communities* (ed. Kerfoot, W. C.) 173–182 (Univ. Press New England, 1980).

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

60. Yoshida, T. et al. Cryptic population dynamics: rapid evolution masks trophic interactions. *PLoS Biol.* **5**, e235 (2007).

61. Vandermeer, J. H. The competitive structure of communities: an experimental approach with protozoa. *Ecology* **50**, 362–371 (1969).

62. Case, T. J. & Bender, E. A. Testing for higher order interactions. *Am. Nat.* **118**, 920–929 (1981).

63. Mickalide, H. & Kuehn, S. Higher-order interaction between species inhibits bacterial invasion of a phototroph–predator microbial community. *Cell Syst.* **9**, 521–533 (2019).

64. Barbosa, M., Fernandes, G. W. & Morris, R. J. Experimental evidence for a hidden network of higher-order interactions in a diverse arthropod community. *Curr. Biol.* **33**, 381–388 (2023).

65. Mayfield, M. M. & Stouffer, D. B. Higher-order interactions capture unexplained complexity in diverse communities. *Nat. Ecol. Evol.* **1**, 0062 (2017).

66. Li, Y. et al. Beyond direct neighbourhood effects: higher-order interactions improve modelling and predicting tree survival and growth. *Nat. Sci. Rev.* **8**, nwaa244 (2021).

67. Lai, H. R., Chong, K. Y., Yee, A. T. K., Mayfield, M. M. & Stouffer, D. B. Non-additive biotic interactions improve predictions of tropical tree growth and impact community size structure. *Ecology* **103**, e03588 (2022).

68. Dyson, F. et al. A meeting with Enrico Fermi. *Nature* **427**, 297–297 (2004).

69. Mayer, J., Khairy, K. & Howard, J. Drawing an elephant with four complex parameters. *Am. J. Phys.* **78**, 648–649 (2010).

70. Tredennick, A. T., Hooker, G., Ellner, S. P. & Adler, P. B. A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology* **102**, e03336 (2021).

71. Aho, K., Derryberry, D. & Peterson, T. Model selection for ecologists: the worldviews of AIC and BIC. *Ecology* **95**, 631–636 (2014).

72. Benincà, E., Ballantine, B., Ellner, S. P. & Huisman, J. Species fluctuations sustained by a cyclic succession at the edge of chaos. *Proc. Natl. Acad. Sci. USA* **112**, 6389–6394 (2015).

73. Chatterjee, S. A new coefficient of correlation. *J. Am. Stat. Assoc.* **116**, 2009–2022 (2021).

74. Rykiel Jr, E. J. Testing ecological models: the meaning of validation. *Ecol. Model.* **90**, 229–244 (1996).

75. Bascompte, J., Jordano, P. & Olesen, J. M. Response to comment on "asymmetric coevolutionary networks facilitate biodiversity maintenance". *Science* **313**, 1887–1887 (2006).

76. Holland, J. N., Okuyama, T. & DeAngelis, D. L. Comment on "asymmetric coevolutionary networks facilitate biodiversity maintenance". *Science* **313**, 1887–1887 (2006).

77. Joly-Smith, E., Wang, Z. J. & Hilfinger, A. Inferring gene regulation dynamics from static snapshots of gene expression variability. *Phys. Rev. E* **104**, 044406 (2021).

78. Wittenstein, T., Leibovich, N. & Hilfinger, A. Quantifying biochemical reaction rates from static population variability within incompletely observed complex networks. *PLoS Comput. Biol.* **18**, e1010183 (2022).

79. Joly-Smith, E. et al. Exploiting fluctuations in gene expression to detect causal interactions between genes. *Elife* **13**, RP92497 (2025).

80. May, R. M. in *It Must be Beautiful: Great Equations of Modern Science* (ed. Farmelo, G.) 212–229 (Granta Books, 2002).

81. Barbier, M., De Mazancourt, C., Loreau, M. & Bunin, G. Fingerprints of high-dimensional coexistence in complex ecosystems. *Phys. Rev. X* **11**, 011009 (2021).

82. Hu, J., Amor, D. R., Barbier, M., Bunin, G. & Gore, J. Emergent phases of ecological diversity and dynamics mapped in microcosms. *Science* **378**, 85–89 (2022).

83. Père, N. M., Terenzi, F. & Werner, B. The dynamic fitness landscape of ageing hematopoiesis through clonal competition. Preprint at *bioRxiv* <https://doi.org/10.1101/2024.04.16.589764> (2024).

84. Remien, C. H., Eckwright, M. J. & Ridenhour, B. J. Structural identifiability of the generalized Lotka–Volterra model for microbiome studies. *R. Soc. Open Sci.* **8**, 201378 (2021).

85. Cho, H., Lewis, A. L., Storey, K. M. & Byrne, H. M. Designing experimental conditions to use the Lotka–Volterra model to infer tumor cell line interaction types. *J. Theor. Biol.* **559**, 111377 (2023).

86. Chen, Y., Angulo, M. T. & Liu, Y.-Y. Revealing complex ecological dynamics via symbolic regression. *BioEssays* **41**, 1900069 (2019).

87. Martin, B. T., Munch, S. B. & Hein, A. M. Reverse-engineering ecological theory from data. *Proc. R. Soc. B* **285**, 20180422 (2018).

88. Cardoso, P. et al. Automated discovery of relationships, models, and principles in ecology. *Front. Ecol. Evol.* **8**, 530135 (2020).

89. Pickett, S. T. in *Long-term Studies in Ecology: Approaches and Alternatives* (ed. Likens, G. E.) 110–135 (Springer, 1989).

90. Wogan, G. O. & Wang, I. J. The value of space-for-time substitution for studying fine-scale microevolutionary processes. *Ecography* **41**, 1456–1468 (2018).

91. Hastings, A. & Powell, T. Chaos in a three-species food chain. *Ecology* **72**, 896–903 (1991).

92. Rogers, T. L., Johnson, B. J. & Munch, S. B. Chaos is not rare in natural ecosystems. *Nat. Ecol. Evol.* **6**, 1105–1111 (2022).

93. Gilpin, M. E. & Ayala, F. J. Global models of growth and competition. *Proc. Natl. Acad. Sci. USA* **70**, 3590–3593 (1973).

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

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

96. Mazzarisi, O. & Smerlak, M. Complexity–stability relationships in competitive disordered dynamical systems. *Phys. Rev. E* **110**, 054403 (2024).

97. d'Ascoli, S., Becker, S., Schwaller, P., Mathis, A. & Kilbertus, N. ODEFormer: symbolic regression of dynamical systems with transformers. In *Proc. Twelfth Int. Conf. on Learning Representations* (2024); <https://openreview.net/forum?id=TzoHLiGVMo>

## Acknowledgements

C.S. and J.M.L. acknowledge funding from National Science Foundation (NSF DEB-2022213). We thank G. Barabas, L. P. Medeiros, S. B. Munch and T. L. Rogers for feedback. We are grateful to members of the Levine laboratory at Princeton for their thoughtful discussion of this paper.

## Author contributions

C.S. and J.M.L. conceived the study. C.S. performed the analysis. C.S. and J.M.L. wrote the paper.

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41559-025-02864-8>.

**Correspondence and requests for materials** should be addressed to Chuliang Song.

**Peer review information** *Nature Ecology & Evolution* thanks Benjamin Martin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2025

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

N/A

Data analysis

All analysis are performed in R 4.4.1. Foundational data import, cleaning, and manipulation are performed using the "tidyverse" suite (v2.0.0), with specific capabilities for reading Excel files provided by "readxl" (v1.4.3). To ensure project portability, file paths are managed with the "here" package (v1.0.1), and all date-time data is systematically handled by "lubridate" (v1.9.4). The graphical visualization process, centered on ggplot2, is significantly enhanced by a collection of specialized packages: "patchwork" (v1.3.0) is used to combine multiple plots into cohesive arrangements, "ggtext" (v0.1.2) enables rich text formatting within graphics, and advanced visual modifications are achieved through "ggh4x" (v0.2.8) for facet and scale customization and "ggblend" (v0.1.0) for sophisticated layer blending. The analytical core of the project relies on the "tidymodels" framework (v1.2.0) for statistical modeling and the "ecoModelOracle" package (v0.1.0; available on <https://github.com/clsong/ecoModelOracle>) for covariance criteria analyses. Finally, computational performance is optimized by implementing parallel processing with "furrr" (v0.3.1) to accelerate demanding calculations.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Empirical data of aquatic invertebrate and the green algae is available from <https://www.nature.com/articles/s41586-019-1857-0>. Empirical dataset of consumer-resource dynamics is available from <https://onlinelibrary.wiley.com/doi/full/10.1111/ele.12291>.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	<a href="#">Human participants, their data, or biological material are not involved in our research.</a>
Reporting on race, ethnicity, or other socially relevant groupings	<a href="#">Human participants, their data, or biological material are not involved in our research.</a>
Population characteristics	<a href="#">Human participants, their data, or biological material are not involved in our research.</a>
Recruitment	<a href="#">Human participants, their data, or biological material are not involved in our research.</a>
Ethics oversight	<a href="#">Human participants, their data, or biological material are not involved in our research.</a>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Our study introduces and applies a novel method, termed the covariance criteria, for the rigorous validation of ecological models against time series data. This approach, rooted in queueing theory, establishes a test based on the covariance relationships between a species' abundance and its model-defined gain and loss rates. We demonstrate the method's utility by applying it to test competing models across three long-standing ecological challenges: the functional form of predator-prey interactions, the role of rapid evolution in population dynamics, and the detection of higher-order interactions.
Research sample	To test our theoretical approach, we use published time series data on species abundances from various ecological systems. The datasets are drawn from several key sources cited in the literature, including a long-term, replicated dataset of an aquatic predator-prey system, a compilation of consumer-resource time series from diverse ecosystems, a long-term dataset of a rocky intertidal community in New Zealand, and a curated subset of single-species time series from the Global Population Dynamics Database.
Sampling strategy	The core of our analysis is the application of the covariance criteria. For a given ecological model and a species' time series, we partition the model's dynamics into terms representing the total gain rate and loss rate. We then calculate the normalized covariances of these inferred rates with the observed species abundance over the time series. The model is considered valid if the normalized covariance of the gain rate with abundance is equal to the normalized covariance of the loss rate with abundance. To assess statistical uncertainty and test this equality, we employ a bootstrapping approach, where we repeatedly resample the time series data with replacement to generate distributions for the two normalized covariance terms.
Data collection	This study involves the analysis of pre-existing, published data. We did not perform new experiments or field observations. The specific data collection protocols for each dataset are detailed in the original publications from which the data were obtained.
Timing and spatial scale	The timing and spatial scale of the data vary across the different studies we analyzed. For example, the rocky intertidal community data were collected monthly over a period of more than 20 years at a single location. The consumer-resource datasets are typically from laboratory microcosm experiments with frequent sampling over many generations. The Global Population Dynamics Database datasets represent a wide variety of sampling frequencies and durations from many distinct locations.

## Data exclusions

Our analysis requires time series of sufficient length and quality to reliably estimate covariance. For the meta-analysis of single-species dynamics, we explicitly used a curated subset of the Global Population Dynamics Database, which excludes time series that are too short or noisy for robust analysis. For other analyses, we relied on datasets previously established in the literature as being of high quality for studying population dynamics.

## Reproducibility

The analyses presented are reproducible. The primary data sources are publicly available and are referenced with citations and, where possible, URLs in the manuscript. All computational methods and statistical analyses were performed using our custom R package, ecoModelOracle, which is publicly available on GitHub.

## Randomization

This is a computational study based on observational and experimental data, so there was no randomization of subjects to experimental treatments. However, randomization is a key component of our statistical procedure. We use bootstrapping, which involves repeatedly drawing random samples (with replacement) from the original time series, to generate distributions of our test statistic and quantify uncertainty.

## Blinding

Not applicable because published data were used and we did not perform experiments or observations.

Did the study involve field work?  Yes  No

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input checked="" type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern
<input checked="" type="checkbox"/>	Plants

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging

## Plants

## Seed stocks

Not applicable because published data were used and we did not perform experiments or observations.

## Novel plant genotypes

Not applicable because published data were used and we did not perform experiments or observations.

## Authentication

Not applicable because published data were used and we did not perform experiments or observations.