Predicting coexistence in experimental ecological communities

Daniel S. Maynard,[1*] Zachary R. Miller,[1] Stefano Allesina[1,2]

[1]Department of Ecology & Evolution, University of Chicago,
1101 E 57th st., Chicago IL 60637, USA
[2]Northwestern Institute on Complex Systems

*Corresponding author e-mail: dmaynard@uchicago.edu

The study of small experimental communities is fundamental to the development of ecology. Yet, for most ecological communities, the number of experiments required to build, model, or analyze the system vastly exceeds what is feasible, rendering these communities experimentally intractable. To address this challenge, we present a statistical approach that uses the results of a limited number of experiments to attempt to infer the outcomes (coexistence and species abundances) of all other possible species assemblages. Using three well-studied experimental systems, we show that this method predicts with high accuracy the results of unobserved experiments. These findings suggest a scalable method for building and exploring large experimental systems, facilitating efforts to study the ecology of diverse natural communities.

Ecologists have long used small experimental systems to explore and test fundamental ecological phenomena: Gause's experiments (*1*) illustrated and popularized the principle of competitive exclusion; chemostat-based systems highlighted the need for eco-evolutionary modeling (*2*); stage-structured populations grown in laboratory conditions offered a clear example of chaotic dynamics (*3*); and yeast communities have provided key insight into ecological tipping points (*4*). However, navigating and studying larger experimental communities has proven exceedingly difficult: a pool of n species could give rise to as many as $2^n - 1$ distinct assemblages—too many to handle for even moderate n—a situation made worse by sensitivity to initial conditions, multistability, complex attractors, and stochasticity.

Experimental ecologists working in speciose systems are thus typically forced to explore a small fraction of the potential assemblages (*5–9*), for example, by culturing each species in isolation, co-culturing all pairs of species (when feasible), and then choosing a few larger combinations of species, often at random (*10*). The hope is that the patterns that emerge in this subset of experiments are representative of the true relationships across all possible communities from the species pool. For example, in one of the largest and most rigorous studies of the relationship between biodiversity and productivity—the "Biodiversity II" experiment, involving 18 plant species—researchers chose 168 out of the $2^{18} - 1 = 262,143$ possible species combinations, tracking their dynamics for decades (*11*). Ideally, we would like to develop principled methods to choose which combinations of species to test experimentally, and to extend the patterns that emerge from these experiments to the full set of communities (*7*).

One common approach for addressing this problem—which has proven successful in disease ecology and other tightly-coupled, low-dimensional systems (*12, 13*)—is based on dynamical modeling. By tracking the abundances of controlled communities across time or space, one parameterizes a dynamical model that can predict the outcomes of unobserved assemblages (*9, 14*). This approach, however, is particularly difficult and expensive, requiring that one first

identify an appropriate model for the dynamics of all $2^n - 1$ assemblages (a non-trivial exercise), and then measure species abundances with sufficient resolution to identify the parameters. More problematically, due to a combination of model misspecification and measurement error, often even the best-fitting models fail to recover known coexistence patterns (*8, 15*), limiting their use as a predictive tool (*9*).

Here we propose an approach for predicting coexistence for all possible assemblages constructed from a pool of n species. Species' abundances are estimated at a single time point in a limited number of experiments, and this information is used to infer the outcomes of unobserved experiments. The approach is agnostic about the dynamics of the system, and provides a scalable method for engineering and studying large experimental ecological communities.

The approach is summarized in Fig. 1: (i) a set of experiments is conducted, with each experiment initialized using different subsets of species from the pool and different initial abundances; (ii) each experimental assemblage is allowed to follow its (unobserved) dynamical trajectory for a sufficiently long time; (iii) the final abundances of the species in each experiment are recorded, in what we term the "endpoint" of the experiment. Crucially, we do not assume that communities reach an equilibrium—experimental endpoints will likely reflect a variety of attractors, including limit cycles or chaos. By conducting multiple replicates of each assemblage under varied initial conditions, one can sample endpoints across an attractor; from this set of points our method attempts to reconstruct the "centroid" of the attractor.

An important experimental consideration is deciding when to sample the abundances of the species. In some cases, there may be a biologically-relevant endpoint (e.g., the end of the growing season for annual plants). In other settings, one might use some inexpensive, non-invasive technique to monitor the state of the community (e.g., optical density, chlorophyll fluorescence, total respiration) until this measure has converged to a stationary distribution. In general, one simply needs to allow enough time to ensure that the dynamics have passed through

any transients and settled into their respective attractors (Supplementary Materials B4).

The simplest, most extensible approach for modeling species' abundances across the experiments is to start from an additive model, in which the average abundance of species i in endpoint k is a linear function of the average abundances of the other species present in endpoint k:

$$\bar{x}_i^{(k)} = \gamma_i + \sum_{j \neq i} \tau_{ij} \bar{x}_j^{(k)} \qquad k = 1, \ldots, 2^n - 1; \quad i \text{ where } \bar{x}_i^{(k)} \neq 0 \ . \tag{1}$$

Here $\bar{x}_i^{(k)}$ is the average abundance of species i in endpoint k (i.e., the "centroid" of the attractor), γ_i models the average abundance of species i when growing alone, and τ_{ij} is the average per-capita effect that members of species j have, if present, on the endpoint abundance of species i. Ecologically, we expect γ_i to be positive for producers—reflecting their carrying capacity—and negative or zero for consumers and predators. More parsimoniously, Eq. 1 can be rewritten as (Materials and Methods):

$$-1 = \sum_j B_{ij} \bar{x}_j^{(k)} \ , \qquad k = 1, \ldots, 2^n - 1; \quad i \text{ where } \bar{x}_i^{(k)} \neq 0 \ , \tag{2}$$

where $B_{ii} = -1/\gamma_i$ and $B_{ij} = \tau_{ij}/\gamma_i$ for $i \neq j$, which we can collect into the $n \times n$ matrix B.

In order to estimate B, we must observe a sufficiently large and diverse set of endpoints. At a minimum, each species must be present in n unique endpoints (not including replicates), and, to estimate each B_{ij} (and B_{ji}), at least one endpoint needs to contain both species i and j together (see Materials and Methods for a precise definition). Under these constraints, we can estimate the matrix B, which amounts to fitting a collection of n hyperplanes (*16*) through the observed endpoints (Fig. 1).

The matrix B encodes all possible endpoints of the system. Thus, armed with an estimate of B, we can determine whether any particular set of species can coexist by taking a matrix inverse (Fig. 1 and Materials and Methods). We have two cases for each predicted endpoint: (i) at least

one species is predicted to have a negative abundance, indicating that the assemblage is unfeasible and cannot coexist; or, (ii) all abundances are positive, indicating that these species can (but need not) coexist. The latter scenario is necessary but not sufficient for coexistence because the endpoint could be dynamically unstable, and thus unreachable experimentally (Materials and Methods).

In order to probe questions about stability, one would need to postulate a dynamical model and estimate additional parameters (e.g., growth rates). For example, as outlined in the Supplementary text, if dynamics were known to be governed by a Generalized Lotka-Volterra (GLV) model, then one could predict the stability of each endpoint and infer whether a given community can be invaded by any other species from the pool (Materials and Methods). In such instances, the coefficients of B are interpretable as rescaled per-capita interaction coefficients—a well-known feature that allows one to approximate the structure of the GLV interaction matrix when growth rates are unknown (*15, 16*). Here, however, we make no assumptions about the dynamics of the system, as our focus is on B as a statistical object that captures the endpoints of the experiments.

Our approach to estimating B and predicting coexistence is sketched in Fig. 1 and described in detail in the Materials and Methods . Although one could estimate B via a series of n separate linear regressions, such an approach ignores the complex error structure of the model. We therefore implement a Bayesian approach, accounting for measurement error in each $x_i^{(k)}$, and for the fact that each endpoint appears in multiple equations (Fig. 1). In doing so, we obtain a posterior distribution for the matrix B and consequently each $\bar{x}_i^{(k)}$, allowing us to compute statistics on the predicted abundances, as well as the probability that a given assemblage coexists.

To demonstrate our approach, we analyze three experimental systems using published data.

First, we consider data from Kuebbing *et al.* (*17*), who conducted growth experiments using old-field plants. They paired four native and four non-native plants, with one plant from each set

belonging to each of the families Asteraceae, Fabaceae, Lamiaceae, and Poaceae. For both sets of species, the system was initialized at 14 out of 15 possible combinations of presence/absence of the species, with each combination replicated 10 times.

Second, we analyze data from Rakowski and Cardinale (*18*), who grew 5-species consumer-resource communities consisting of five species of green algae and one of two herbivorous species from the family Daphniidae (*Ceriodaphnia dubia* and *Daphnia pulex*). Here we focus on the four algal species that survived in a sufficiently large number of endpoints to test our model: *Chlorella sorokiniana*, *Scenedesmus acuminatus*, *Monoraphidium minutum* and *Monoraphidium arcuatum*. The algae were grown in all four-species combinations with high replication, and an herbivore was added in two-thirds of the replicates. The communities were incubated for 28 days, during which time each assemblage collapsed to a subset of the initial pool.

Third, we turn to time-series data published by Pennekamp *et al.* (*19*), who cultured six ciliates species growing on bacterized medium, capturing 53 of the 63 possible assemblages. We omit one species that declined in abundance across the time series (*Tetrahymena thermophila*), leading to the five-species subsystem used here: *Colpidium striatum*, *Dexiostoma campylum*, *Loxocephalus* sp., *Paramecium caudatum*, and *Spirostomum teres*. Each experimental assemblage was replicated twice and tracked for 57 days (*20*), and experiments were repeated at six different temperatures ranging from 15 to 25° C.

These three systems comprise 10 distinct experimental scenarios: a native and nonnative plant community, two different herbivore-algae sub-communities, and six protist temperature conditions. Each of these subsystems contains a sufficient number of endpoints to fit a model and test the accuracy on out-of-fit assemblages, using a "jackknife" leave-one-out approach (Supplementary Materials A). By applying our method to each of these data sets, we obtain a posterior distribution for the average abundance of each species in each community, as well as a likelihood of coexistence for the unobserved and out-of-fit assemblages.

As illustrated in Fig. 2, our method predicts with high accuracy the endpoints of each of the systems, with the medians of the posterior distributions closely tracking the median observed abundances. Not only does our method achieve high quality-of-fit for within-sample endpoints (Fig. 2a, c, e); it is also able to estimate with high accuracy the out-of-fit endpoints (Fig. 2b, d, f) and approximate the observed range in abundances for each species in each community.

Although the method does not allow for a mechanistic interpretation of B—which we view primarily as a means to an end—this matrix can illustrate the consistency of this approach while hinting at important biological processes. For example, we find that the effects of one plant on another are phylogenetically conserved across the native and non-native plant data sets (Fig. 3a, Supplementary Materials A2). Similarly, in the herbivore-algae system, our results demonstrate that the identity of the herbivore has little effect on interactions among the algae, suggesting weak or non-existent herbivore-mediated higher-order effects (Fig. 3b, Supplementary Materials A3). In the protist system, we can compare the structure of B across the temperature gradient, revealing consistent changes (in magnitude and direction) of the B_{ij} across temperatures, and suggesting a mix of positive and negative effects on the endpoint abundance (Fig. 4, Supplementary Materials A4). Due to statistical nature of our method, however, any strict dynamical or mechanistic interpretation of these pairwise relationships would require further testing. Nonetheless, these patterns highlight that this approach yields results which are ecologically consistent across communities.

We have successfully applied our method to three experimental systems, and although our model assumes the endpoint structure is approximately linear and additive, we crucially do not believe that the dynamics of these systems are strictly linear or pairwise. We completely neglect functional responses, behavioral changes, spatial effects, resource depletion, etc.—all of which have been documented in these systems to varying degrees (*21–25*). However, by foregoing dynamics, we gain tractability, leading to improved predictive power. Indeed, in the protist sys-

tem, our estimate of B compares favorably to what we obtain by fitting a GLV dynamical model to the full time-series (Supplementary Materials B), yet our method predicts the experimental endpoints with higher accuracy while requiring a fraction of the data. Such results demonstrate that the endpoints of an ecological system can be approximately additive even if the dynamics are not.

To explore the generality of this finding, we fit simulated data under a variety of non-linear, non-equilibrium, and non-pairwise dynamical models (Supplementary Materials C). In all cases, our approach recovers the endpoints of the systems quite accurately, not only when approximating the abundances of "observed" communities, but also in predicting combinations for which coexistence is impossible. These simulation results demonstrate the generality and flexibility of Eqns. 1-2, further highlighting that nonlinear dynamical systems can contain an approximately additive endpoint structure.

Nevertheless, there are undoubtedly ecological systems where this approach will fail. Most obviously, this approach cannot be applied to systems where coexistence is rare, as this would provide too few equations to estimate all the coefficients of B. In such settings, one can prune the system to the subset of species that do exhibit sufficiently widespread coexistence, as was done here for both the protist and herbivore-algae systems. This approach should also perform poorly when there is true multi-stability, in which case the predicted abundances will reflect an average across distinct attractors.

More generally, the predictive power of this method will decline when the direction of the interaction between two species varies across communities (e.g., a competitive interaction becoming mutualistic when another species is added). Indeed, in the simulations, our method struggled most with systems containing strong higher-order interactions (Supplementary Materials C6), in which the effect of one species on another was mediated by the abundance of a third species. Yet a benefit of focusing on endpoints rather than dynamics is that these nonlin-

ear relationships can easily be incorporated into our method by including additional terms in Eqn. 1, such as by summing over all pairs of species, $\sum_{j,m} \tau_{ijm} x_j^{(k)} x_m^{(k)}$—analogous to incorporating interactions into a standard linear regression model. Doing so allows one to compare various model formulations and test goodness-of-fit on out-of-fit experiments, so as to identify a model that best describes the structure of endpoints. In this way, exploring more complex models does not require more complex experimental designs or substantially more data, as are often necessary when parameterizing non-pairwise dynamical models (*26, 27*).

Lastly, because multispecies endpoints contribute to multiple equations (Fig. 1), large communities contain more information about B than do small communities. Leveraging this feature, our method provides an alternative approach to experimental design which scales linearly with the size of the species pool. We show (Supplementary Materials D) that the common approach of growing species in monoculture and in all pairwise combinations (*7–9, 15, 28*) can be highly inefficient, requiring upwards of $n^2/2$ experiments while providing poor predictive power in the presence of noise. In contrast, by combining species-poor and species-rich experiments—such as all monocultures and all leave-one-out assemblages—one more efficiently 'anchors' the slopes of the fitted hyperplanes, providing high predictive power while requiring approximately $2n$ experiments.

These results suggest a simple, iterative method for building large experimental communities: one first conducts a small number of experiments, sufficient to obtain a draft estimate of B, which is then used to identify subsequent assemblages where species are predicted to coexist with high probability. By iteratively updating B and testing these predictions on new, out-of-fit experiments, one can obtain a parameterization that robustly predicts the full set of coexistence patterns. This approach can help experimentalists navigate the enormous space of possible assemblages by providing an efficient way to explore, build, and quantify large experimental systems.

References and Notes

1. G. F. Gause, *The struggle for existence* (The Williams & Wilkins Company, Baltimore, 1934).

2. T. Yoshida, L. E. Jones, S. P. Ellner, G. F. Fussmann, N. G. Hairston Jr, *Nature* **424**, 303 (2003).

3. R. F. Costantino, R. Desharnais, J. M. Cushing, B. Dennis, *Science* **275**, 389 (1997).

4. L. Dai, D. Vorselen, K. S. Korolev, J. Gore, *Science* **336**, 1175 (2012).

5. M. W. Cadotte, *Proc. Natl. Acad. Sci.* **110**, 8996 (2013).

6. D. a. Heemsbergen, *et al.*, *Science* **306**, 1019 (2004).

7. J. Friedman, L. M. Higgins, J. Gore, *Nat. Ecol. Evol.* **1**, 1 (2017).

8. N. J. B. Kraft, O. Godoy, J. M. Levine, *Proc. Natl. Acad. Sci.* **112**, 797 (2015).

9. C. F. Dormann, S. H. Roxburgh, *Proc. R. Soc. B Biol. Sci.* **272**, 1279 (2005).

10. T. Bell, J. A. Newman, B. W. Silverman, S. L. Turner, A. K. Lilley, *Nature* **436**, 1157 (2005).

11. D. Tilman, *et al.*, *Science* **294**, 843 (2001). A total of 168 plots were seeded in 1994 and again in 1995 with either one, two, four, eight or sixteen plant species out of a pool of 18. By 2015, 122 (72%) of the plots had lost at least one of the originally-planted species; in fact, all of the plots containing either 8 or 16 plants had lost at least one of the initial species.

12. M. J. Smith, *et al.*, *Proc. Natl. Acad. Sci.* **106**, 7905 (2009).

13. M. Pascual, X. Rodo, S. P. Ellner, R. Colwell, M. J. Bouma, *Science* **289**, 1766 (2000).

14. E. T. Borer, P. R. Hosseini, E. W. Seabloom, A. P. Dobson, *Proc. Natl. Acad. Sci.* **104**, 5473 (2007).

15. F. Carrara, A. Giometto, M. Seymour, A. Rinaldo, F. Altermatt, *Methods in Ecology and Evolution* **6**, 895 (2015).

16. Y. Xiao, *et al.*, *Nature communications* **8**, 2042 (2017).

17. S. E. Kuebbing, A. T. Classen, N. J. Sanders, D. Simberloff, *New Phytologist* **208**, 727 (2015).

18. C. Rakowski, B. J. Cardinale, *Oikos* **125**, 1627 (2016).

19. F. Pennekamp, *et al.*, *Nature* **563**, 109 (2018).

20. F. Pennekamp, *et al.*, *PloS one* **12**, e0176682 (2017).

21. J. W. Fox, *Oikos* **116**, 189 (2007).

22. K. M. Crawford, T. M. Knight, *Oecologia* **183**, 211 (2017).

23. O. S. Beveridge, O. L. Petchey, S. Humphries, *J. Anim. Ecol.* **79**, 1324 (2010).

24. D. E. Golberg, *Ecology* **68**, 1211 (1987).

25. O. Sarnelle, A. E. Wilson, *Ecology* **89**, 1723 (2008).

26. A. D. Letten, D. B. Stouffer, *Ecol. Lett.* **22**, 423 (2019).

27. J. M. Levine, J. Bascompte, P. B. Adler, S. Allesina, *Nature* **546**, 56 (2017).

28. D. S. Maynard, *et al.*, *Nat. Ecol. Evol.* **1**, 0156 (2017).

ACKNOWLEDGMENTS

We thank S. Kuebbing, C. Rakowski, F. Pennekamp, and O. Petchey for making their data available and for comments on earlier drafts of this manuscript. We thank C. Serván, M. Pascual, J. Bergelson, E. Baskerville, G. Barabás, and E. Friedlander for assistance and suggestions throughout this study.

Author contributions. DSM, ZRM, and SA conceived of this study and developed the methods. DSM collected and analyzed the data, and wrote the supplemental data analysis. SA and ZRM wrote the supplemental materials and methods and implemented the simulations. All authors contributed to the writing of the manuscript.

Competing interests. The authors declare no competing interests.

Data and materials availability. All data and code needed to reproduce these results are available at http://github.com/dsmaynard/endpoints.

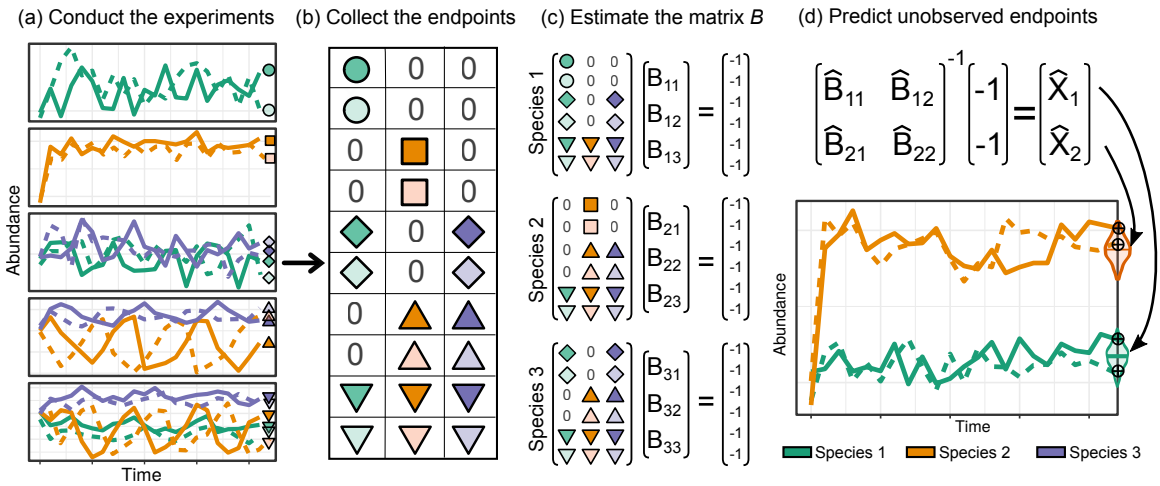

Figure 1: **Predicting unobserved experimental outcomes**. (a) A set of experiments is conducted, with each assemblage varying in initial composition and initial abundances. Colors denote different species, and line types each of two replicates. The experiments are run for a sufficiently long time, at which point a single snapshot is taken of each species' abundance in each replicate. (b) These abundances are collected into a matrix, with columns corresponding to species and rows corresponding to endpoints. Here, shapes represent species' endpoint abundances for each of two replicates across five different assemblages. (c) The matrix B, which defines the full set of endpoints, is estimated by fitting a separate hyperplane through each species' set of endpoints. (d) To predict the endpoint of an unobserved assemblage—here, the combination of species 1 and 2—we take the corresponding sub-matrix of B, invert it, and compute the negative row sums. To analyze the three empirical datasets, we implement a Bayesian approach to estimating B, resulting in a posterior distribution for the unobserved endpoint abundances.

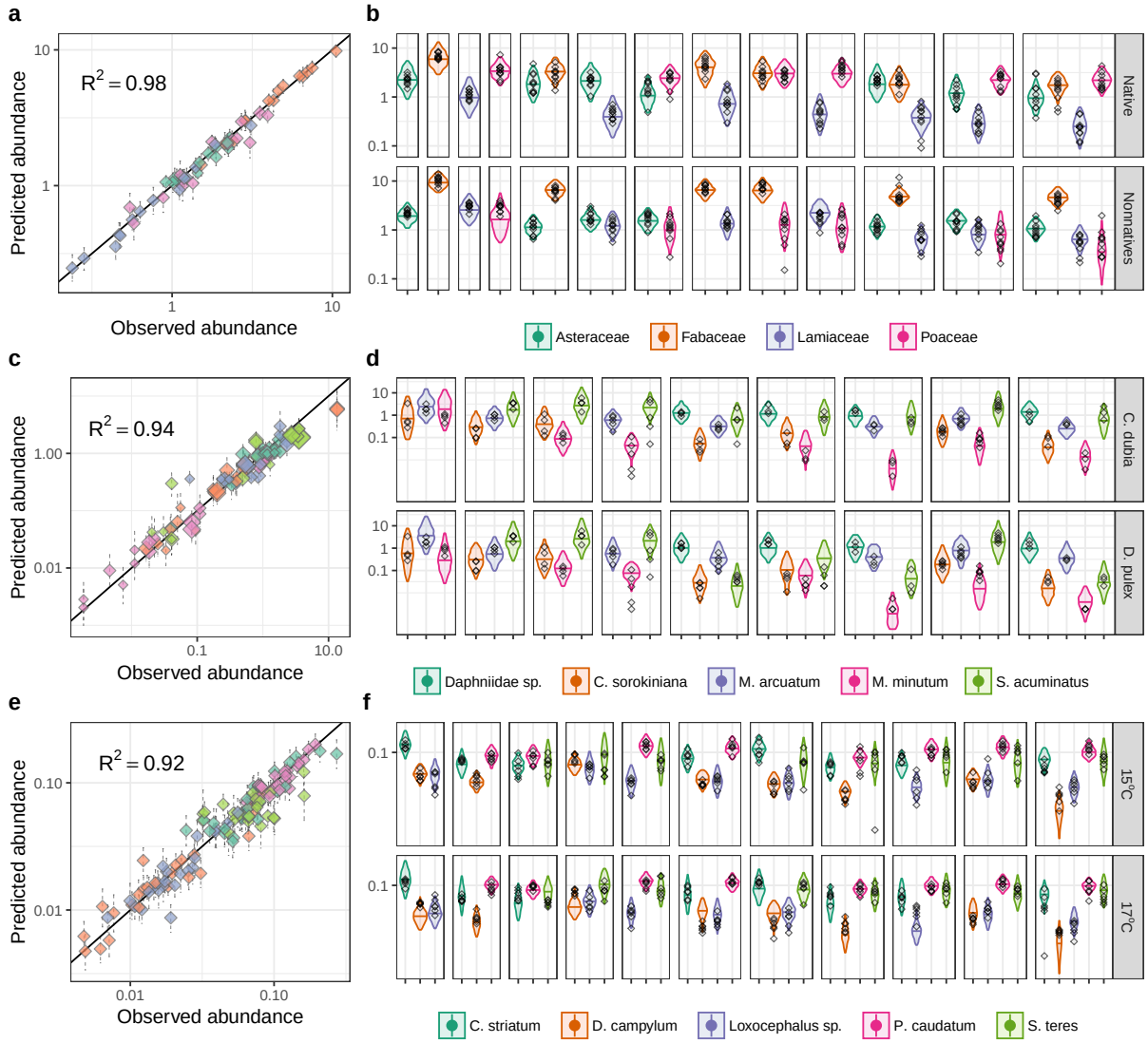

Figure 2: **Results for three experimental systems**. (a, b) The old-field plant system comprising a native and non-native community; (c, d) the herbivore-algae system comprising two different Daphniidae species; (e, f) the ciliated protist system, assayed under 6 different temperature conditions, with 15° and 17°C shown here. (a, c, e) The observed vs. predicted abundances for each experimental system, with the vertical error bars denoting the 95% prediction interval of the posterior distribution. (b, d, f) The out-of-fit predicted abundances for each assemblage, with horizontal bars indicating the median predicted values, and violin plots showing the 95% prediction intervals for the posterior distribution. Black points show the measured abundances at the end of each experiment. Note that only communities containing more than two species are shown in d and f; see Figs. S8-S11 for the complete results.

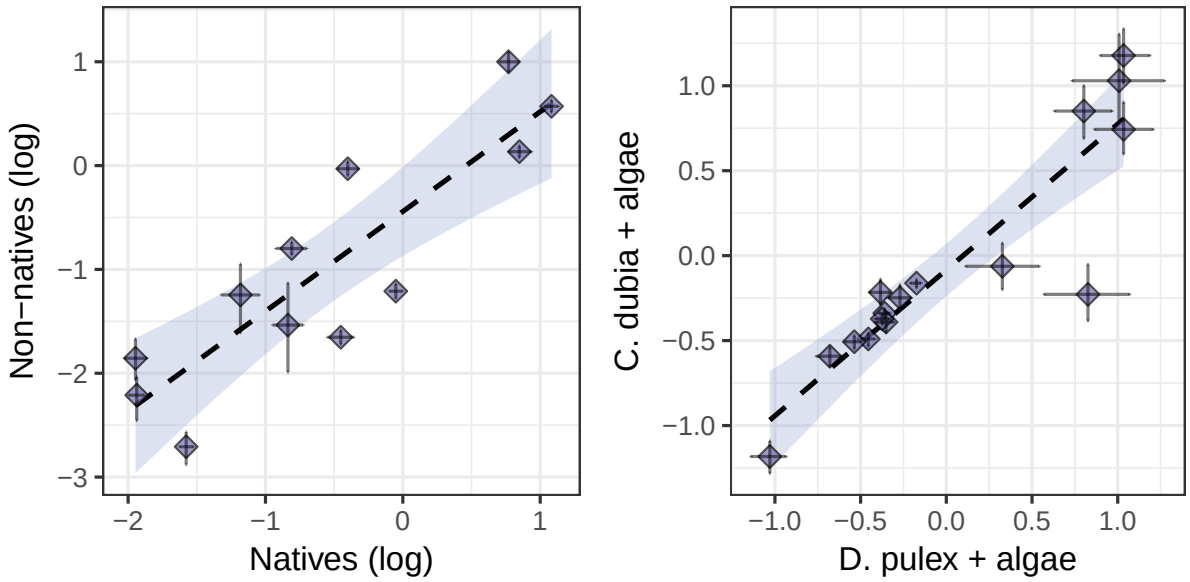

Figure 3: **Comparison of the matrix B across the plant and herbivore-algae datasets**. (a) The coefficients for the native vs. non-native plant systems, aligned by family into Asteraceae, Fabaceae, Lamiaceae, and Poaceae, reveal that the effects of one family of species on the others are conserved across the two systems. (b) The algae-by-algae coefficients for the herbivore-algae system, estimated for the communities containing *D. Pulex* vs. those containing *C. dubia*. The coefficients are generally consistent across both systems, suggesting minimal herbivore-induced changes (potentially, higher-order effects), albeit with the effect of *C. sorokiniana* on *M. arcuatum* estimated to be negative in the presence of *C. dubia* but positive in the presence of *D. pulex*.

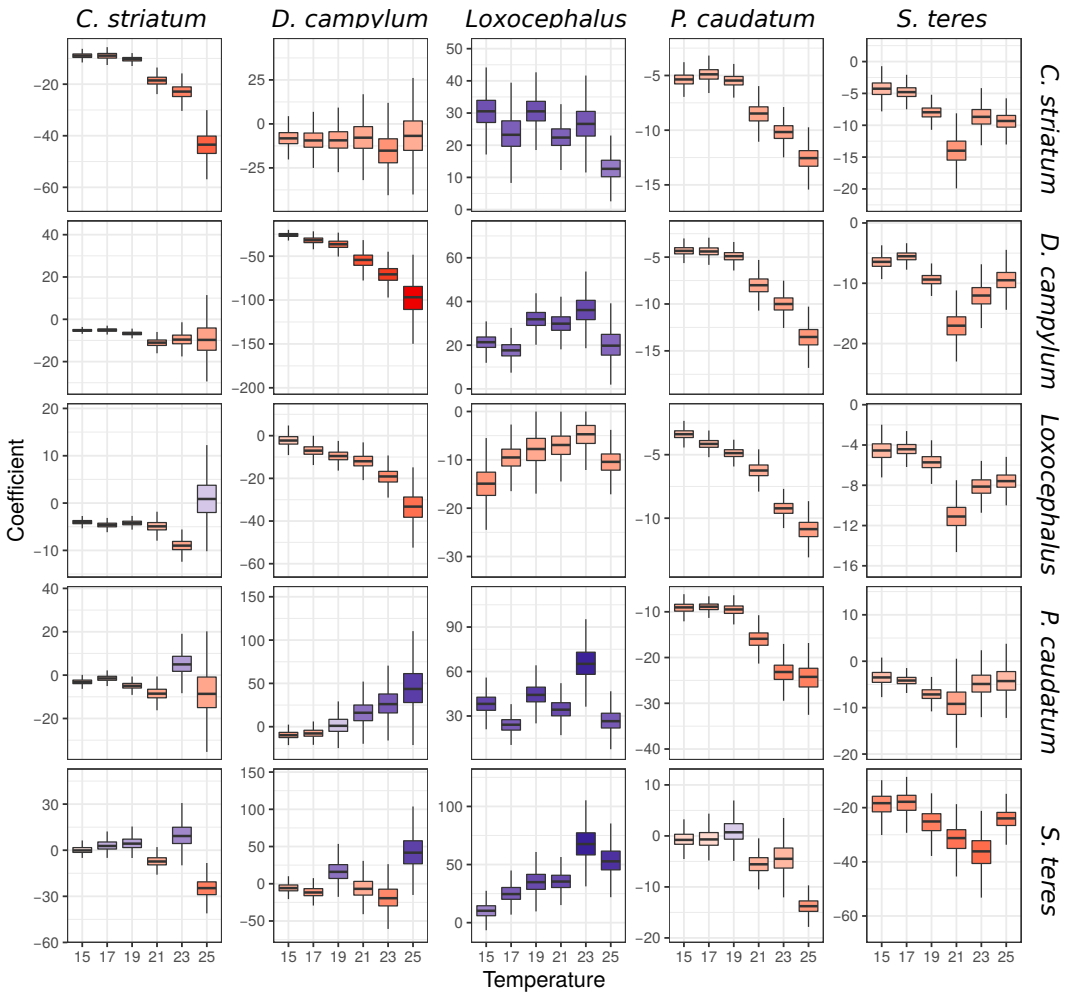

Figure 4: **Comparison of the matrix B in the protist system across the six temperatures**. The boxplots report the medians and interquartile ranges for the posterior distribution of each coefficient B_{ij}, with blue indicating positive effects and red indicating negative. Our method reveals ecologically consistent changes in these coefficients across the temperature gradient.