Linking Pattern With Process in (Meta)community ecology: Challenges and Opportunities

Mathew A. Leibold, Department of Biology, University of Florida, Gainesville, Florida 32661, USA and Institut Natura e Teoria en Pireneus, Surba, France. ORCID: 0000-0003-3954-3187. mleibold@ufl.edu

Matthieu Barbier, CIRAD, UMR PHIM, 34090 Montpellier, France. PHIM Plant Health Institute, University of Montpellier, CIRAD, INRAE, Institut Agro, IRD, 34090 Montpellier, France. ORCID: 0000-0002-0669-8927. contact@mrcbarbier.org

Leonora S. Bittleston, Department of Biological Sciences, Boise State University, Boise, Idaho 83725. ORCID: 0000-0003-4007-5405. leonorabittleston@boisestate.edu

Adam T. Clark, Department of Biology, University of Graz, Holteigasse 6, 8010, Graz, Austria. ORCID: 0000-0002-8843-3278. adam.clark@uni-graz.at

Catalina Cuellar-Gempeler, Department of Biological Sciences, California Polytechnic University, Humboldt. 1st Harpst Street, Arcata, CA. ORCID: 0000-0002-2634-9599. ccg@humboldt.edu

Rafael D'Andrea, Department of Ecology and Evolution, Stony Brook University, Stony Brook, 11790, USA. ORCID: 0000-0001-9687-6178. rafael.dandrea@stonybrook.edu

Veronica F. Frans, Hopkins Marine Station, Stanford University, Pacific Grove, CA, USA. ORCID: 0000-0002-5634-3956. verofrans@gmail.com

Gabriel Khattar, Department of Biology, Concordia University, Montreal, Quebec, Canada. 0000-0002-1703-4210. gabriel.khattar@concordia.ca

Zachary R. Miller, Department of Earth and Planetary Sciences, Yale University, New Haven CT 06511, USA. ORCID: 0000-0001-9762-8301. zachary.miller@yale.edu

Pedro R. Peres-Neto, Department of Biology, Concordia University, Montreal, Quebec, Canada. ORCID: 0000-0002-5629-8067. pedro.peres-neto@concordia.ca

Nathan I. Wisnoski, Department of Biological Sciences, Mississippi State University, Mississippi State, MS 39762, USA. ORCID: 0000-0002-2929-5231.
nathan.wisnoski@msstate.edu

Running Title: "Pattern to Process and Back"

Keywords: Predictive models, Inverse Inference, Joint Species Distribution Models, Disordered Systems Models, Metacommunities, Community Entanglement.

Dedication: We dedicate this paper to Gabriel Khattar who passed away during the review process. He was an outstanding and promising young community ecologist who we all value tremendously as a colleague and friend.

Statement of authorship: This paper is the result of a workshop funded by NSF award to MAL. All co-authors actively participated in discussions during the workshop and remote discussions after the workshop. MAL and MB co-wrote the initial draft of the paper and all co-authors provided extensive edits and comments leading up to the final draft.

Data statement: This manuscript has no data.

Abstract (159 words)

Ecological communities, and especially metacommunities, are complex and dynamic entities. Resolving the processes and mechanisms that shape these systems remains a central challenge in ecology. This challenge is compounded by the increasing entanglement of mechanisms, processes, and emergent patterns of biodiversity as scales of space, time, and biological organization expand. Here, we define and contextualize key issues, describe recent progress, and identify remaining challenges in interpreting basic metacommunity data and using predictive models to link processes to patterns and back.

We identify two contrasting modeling strategies for complex metacommunities, "top-down" and "bottom-up", and we consider how they guide different approaches to pattern-to-process inference. We find substantial progress in connecting pattern and process through improved data repeatability and scaling, enhanced analytical tools to quantify patterns, and increasingly sophisticated theoretical models that address ecological complexity. However, accurately matching observable patterns with process-oriented theory remains a persistent challenge. Finally, we identify potential pipelines connecting process and pattern and highlight areas for future progress.

Key words: Communities, metacommunities, patterns, metrics, processes, mechanisms, disordered systems models, joint species distribution models.

Introduction:

Ecological communities are shaped by a complex interplay of a limited number of basic ecological processes, or "forces", that influence species establishment, growth, and persistence. While the number of fundamental **processes** is considered small (Vellend 2010, 2015), their combinations generate a wide variety of recognized **mechanisms** that organize community dynamics and structure (a glossary of terms in bold is provided in Table S1). This complexity is especially pronounced in systems with many species and when local communities interact across space and time as metacommunities.

The interplay among ecological processes (see below) gives rise to the patterns observed in metacommunity data. Recent advances in data collection and analytical methods have significantly improved our ability to detect, describe, and quantify metacommunity patterns. Yet, inferring the underlying processes and their relative importance from these patterns remains a major challenge in (meta)community ecology (Schaffer 1981; Sanderson and Pimm 2015; Leibold et al. 2022).

Vellend (2010, 2015), drawing an analogy with evolutionary theory, proposed that a handful of basic processes shape community assembly and long-term dynamics. As in Thompson et al. (2020), we reframe Vellend's original processes to better align with contemporary ecological theory:

a) **Density-independent selection**: Species growth and persistence depend on their responses to environmental conditions, often (though not exclusively) abiotic. No feedback from the biota to these environmental factors over time scales of interest is assumed.

b) **Density-dependent selection**: Biotic interactions with conspecifics or heterospecifics, including direct interactions and indirect effects mediated by other species that can provide feedback mechanisms that are not accounted for by density-independent selection.

c) **Dispersal**: Movement of organisms among local communities. Dispersal limitation restricts species to subsets of suitable sites, while dispersal excess allows persistence in otherwise unsuitable habitats. This could also include various forms of dormancy, which can be thought of as temporal dispersal. An important consequence of adding dispersal in metacommunity ecology is that it implies that spatial connectivity can be an important component driving community variation.

d) **Novelty and trait diversification**: Includes speciation, evolutionary change, and species introductions due to biogeographic shifts or anthropogenic influence (e.g., introductions). This component remains underexplored in metacommunity studies (but see Borregaard et al. 2014; Germain et al. 2021; Leibold et al. 2023).

e) **Stochasticity**: While stochasticity permeates all the above processes, we specifically highlight demographic stochasticity (random birth and death events, especially in small populations) and temporal environmental variation. These sources of variability are often difficult to distinguish from measurement error (Shoemaker et al. 2020) and may therefore be conflated in empirical analyses. Treating these sources of stochasticity as variance components provides a practical way to account for their combined effects when analyzing metacommunity data without requiring explicit mechanistic models for each source of variability.

In nature, these processes interact to generate various **mechanisms** (e.g., resource partitioning, species sorting, mass effects, trophic cascades) with distinct effects on community patterns. **Models** often assume specific relationships among processes (e.g., trade-offs or context dependence) to represent ecological mechanisms, typically in simplified systems involving few species. Classic models like Lotka–Volterra (Lotka 1925; Volterra 1927) or resource competition models (MacArthur 1974; Tilman 1982) illustrate how abiotic trade-offs and species interactions shape coexistence. Experimental validation of such models has been demonstrated in simple plant, animal, and microbial communities (e.g., Gause 1932, 1934; Crombie 1944; Vandermeer 1965; Tilman 1980; see Kneitel and Chase 2004).

However, as the number of species increases, the potential combinations of **parameters** grow exponentially (and sometimes even factorially!), and multiple mechanisms often operate simultaneously. Most natural patterns emerge not from isolated species interactions but from the intertwined web of biotic interactions and environmental effects, i.e., Darwin's "entangled bank" (Darwin 1859; Schaffer 1981; Kéfi et al. 2016).

While we understand how specific processes can generate different distributional patterns in metacommunities, reliably inferring the underlying processes and mechanisms from observed patterns remains a major challenge. Multiple distinct process-based models can predict similar patterns (e.g., Barbier et al. 2018). This many-to-one mapping is both an opportunity and a challenge for ecology. It means that our models may effectively forecast responses to change (e.g., under climate change scenarios) without explaining the causal mechanisms involved. Such predictive capacity can be highly valuable for ecological applications, including policy and management.

Yet, as René Thom noted, "To predict is not to explain" (Thom et al. 2016). As environmental changes push systems beyond the conditions that informed past predictions, explanatory understanding, along with predictive capacity, becomes essential. Identifying the processes underlying observed patterns is therefore critical for addressing contemporary environmental challenges, strengthening empirical ecology, and advancing ecological theory. Our goal here is to evaluate the extent to which current methods link pattern and process and, in doing so, improve our understanding of the entangled web of ecological dynamics in biodiverse communities and metacommunities.

We structure this essay as follows:

1. We begin by defining “patterns” in metacommunities, contrasting those found in individual communities with those specific to metacommunities.
2. We then explore “processes” along with associated concepts such as mechanisms and models. We contrast two modeling approaches: a reductionist, “bottom-up” strategy and an emergent, “top-down” approach inspired by statistical mechanics, while recognizing that other conceptual and modeling approaches are also possible.
3. We evaluate whether basic ecological processes can be robustly linked to patterns to test or validate models and to make predictions.
4. We consider the inverse problem: how observed patterns can be used to infer underlying processes and explain metacommunity dynamics.
5. Finally, we contextualize these discussions within current trends in community and metacommunity ecology, identifying opportunities to overcome outstanding challenges.

Taken together, these points offer a roadmap for integrating process-based and pattern-based approaches in community and metacommunity ecology. Despite the difficulty of linking pattern and process, we argue that there are strong reasons for optimism. Advances in data generation, modeling approaches (theoretical, statistical, and computational), and theoretical frameworks position community and metacommunity ecology for continued and exciting progress.

Data and Patterns

We begin by clarifying what we mean by "**data**" in the context of metacommunities. Data are direct observations of community variation across localities within a region and are typically organized as site-by-species matrices. These matrices may represent presence/absence, abundance, relative abundance, biomass, or other ways to describe and quantify community composition and structure. Data can originate from natural systems, experimental studies, or even simulations of ecological models. From these matrices, we derive "patterns": simplified, often quantitative, summaries of ecological structure or variation.

Patterns range from local community attributes (e.g., species richness, diversity indices, species abundance distributions, and compositional turnover) to aggregated summaries across sites (e.g., beta-diversity or variability and synchrony in temporal fluctuations). Comparisons among sites often incorporate ancillary variables like environmental gradients, spatial structure, or spatial isolation.

Metacommunity-level patterns consider spatially structured assemblages of multiple, interconnected communities. This broader focus enables exploration of how spatial dynamics such as dispersal interact with other structuring forces, such as internal patch dynamics and environmental differences among patches. For example, Leibold and Mikkelson (2002) proposed the "Elements of Metacommunity Structure" to characterize coexistence patterns in spatially explicit metacommunities. Similarly, Cottenie (2005) applied variation partitioning methods (Borcard et al. 1992; Peres-Neto et al. 2006) to separate spatial and environmental components underlying community variation, aiming to diagnose dominant ecological processes.

Together, such methods provide a diverse set of statistical models and corresponding pattern metrics that characterize communities and metacommunities (e.g., Ovaskainen et al. 2019; Thompson et al. 2020, Guzman et al. 2022). Each metric serves as a potential clue about the processes and mechanisms shaping communities, though few are diagnostic on their own.

For example, researchers frequently examine how observed patterns deviate from null models, which are baseline expectations assuming the absence of particular ecological processes by randomization. While null models can become complex and their utility contentious (e.g., Peres-Neto et al. 2001), most patterns deviate significantly from null expectations in at least some systems (Gotelli and McCabe 2002; Cottenie 2005).

There are a number of increasingly complex statistical models for analyzing metacommunity structure, including beta-diversity partitioning (Si et al. 2017), network models (e.g. Borthagarai et al. 2014), and Joint Species Distribution Models (JSDMs; Ovaskainen et al. 2017) and other forms of Generalized Latent Variable Models (GLMVMs, Skrondal and Rabe-Hesketh 2004). JSDMs have become a good example of models that evaluate the major processes identified by Vellend (2010) described above, as they allow for explicit statistical modeling of co-occurrence structures that account for environmental and spatial variation. They also provide a powerful framework for inferring latent (unmeasured) influences beyond measured environmental and spatial variables or species trait variation, including potential effects of biotic interactions (although these cannot generally be separated from unmeasured environmental and spatial factors; see Blanchet et al. 2020 for a discussion).

However, such analyses alone do not identify the specific processes underlying metacommunity structure, as multiple models based on different mechanisms can produce convergent predictions. To improve inference, ecologists increasingly use cross-validation across multiple patterns (Holling and Allen 2002; Yanco et al. 2020). For instance, May et al. (2015) showed that while a neutral model (similar to Hubbell 2001) could individually match several empirical patterns from Barro Colorado Island, it could not do so with consistent parameter estimates. As a result, the neutral model can be rejected because it fails to jointly explain multiple patterns, even though it appears plausible when each pattern is considered in isolation.

This strategy of testing models against suites of independent metrics helps identify which mechanisms are more plausible. While not definitive, such multi-pattern approaches (e.g., Ovaskainen et al. 2019; Thompson et al. 2020; Guzman et al. 2021) represent a practical path toward more robust ecological inference.

Processes, Models, and Mechanisms

While patterns arise from data, we define "processes" as the underlying forces and interactions that generate those patterns. Unlike patterns, processes are not easily quantified, and are instead inferred, often through theory-informed interpretations of pattern analysis or through **mechanistic modeling**. This distinction between what is observed and what is hypothesized to underlie the observations has long posed a challenge in ecology, and definitions of processes vary widely in the literature (see Section 2 in the Supplementary Information).

Although Vellend’s five-process framework is comprehensive (Vellend 2010, 2015), the ways in which these processes can interact to shape ecological patterns remain unclear. In the short-term, these processes operate in a quasi-linear fashion in theoretical frameworks based on e.g., metacommunity Lotka–Volterra models (Gravel et al. 2016) or colonization–extinction models (Leibold et al. 2022, Jackson et al. 2025). Yet, over longer timescales or under more complex ecological dynamics (e.g., models with higher-order interactions), the effects of these processes become entangled in ways that obscure clear causal links. Understanding this entanglement is central to resolving how pattern and process are related in community and metacommunity ecology.

Historically, ecologists have used mechanistic theories to explain long-term patterns in ecological communities. One common strategy involves identifying a plausible mechanism, such as interspecific competition, predation, and/or mutualism, and modeling it with systems of differential (or difference) equations. For example, Lotka (1925) and Volterra (1927) modeled interspecific competition and predator–prey dynamics. These foundational models were empirically validated in microcosm experiments by Gause (1932, 1934) and others (e.g., Crombie 1945, 1946).

Such experiments were instrumental in testing theories involving density-dependent biotic interactions; however, they often overlooked density-independent processes such as environmental filtering (e.g., Choler et al. 2001) or the role of spatial dynamics and dispersal (Amarasekare 2003; Peres-Neto et al. 2012, but see Huffaker 1958).

Alternative models have focused on community patterns without assuming strong species interactions. These include spatial models assuming weak or no interspecific interactions (MacArthur and Wilson 1967; Connor and Simberloff 1979; Hart and Newman 2014) or neutral models in which all individuals are ecologically equivalent (Hubbell 2001). More comprehensive models that integrate species interactions, environmental filtering, and spatial dynamics are still relatively rare and often rely on complex simulations that are difficult to interpret, especially in species-rich or spatially structured systems.

Bottom-Up versus Top-Down Modeling Perspectives

As more species are included, model complexity increases dramatically (see Appendix 1). A traditional bottom-up approach to tackling this complexity is to model small sets of species using mechanistic **modules** (Holt and Hochberg 1999). These modules allow for detailed analysis of interactions, including indirect and higher-order effects, often in relation to environmental or spatial contexts (e.g., Tilman 1982; Holt et al. 1994; Leibold 1996). However, such models become analytically intractable with more than a few species, and even three-species systems can be difficult to fully analyze (e.g., Ranjan et al. 2024). While they serve as an informative starting point, these modules inevitably oversimplify the dynamics of natural communities, which involve complex webs of direct and indirect interactions (Schaffer 1981; Kéfi et al. 2016).

An alternative is a top-down approach that seeks to explain patterns at the aggregate level using simplified assumptions about species interactions. As a classic example, May (1970, 1972) used random matrix theory to analyze stability in large communities. He assumed that interaction coefficients among species were randomly distributed, and showed that community stability occurs when the number of species (S), the connectance (c), and the standard deviation of interaction strengths (σ) obey the inequality:

$$\sigma \sqrt{c\,(S-1)} \,<\, m$$

where m is the mean intraspecific self-limitation term. According to this inequality, communities with too many strongly connected species interactions as compared with self-limitation are unlikely to be stable.

These "disordered systems" models rely on minimal assumptions and treat interaction networks as random objects, summarized by statistical moments (means, variances). However, many ecologists are uncomfortable with their oversimplified assumptions. For instance, Yodzis (1981) noted that trophic structure was a missing component in May’s original formulation.

To address this, researchers have developed "partially structured" models (Ahmadian et al. 2015; Barbier et al. 2018; Carugno et al. 2022; Servan et al. 2025) that incorporate limited structure into interaction matrices such as distinct subgroups (guilds), body size scaling, trophic levels, or evolutionary relationships. Reviews by Akjouj et al. (2024) and Cui et al. (2024) explore how such partial structure can be introduced into community matrices to increase realism without resorting to fully detailed complexity.

A clear question arises: how much structure should we impose? While including trophic structure might seem essential, increasing structure (e.g., spatial correlations or functional traits) eventually risks reverting to highly constrained, bottom-up models. If over-structured, these models may become as complex and difficult to analyze as the bottom-up models they were intended to complement!

In sum, we can approach ecological modeling from two strategic directions (Figure 1). The bottom-up approach starts with simple population models and incrementally adds complexity. Predicted patterns are closely linked to mechanisms, but this approach is limited to systems involving small sets of species. The top-down approach discussed here, inspired by statistical mechanics, uses randomized or partially structured models (e.g., interaction matrices) to make robust but coarse-grained predictions in highly biodiverse systems. By focusing on aggregate system properties rather than detailed interactions among all components, statistical mechanics provides a natural framework for deriving general expectations in complex, high-dimensional ecological systems. Each strategy has limitations: bottom-up models lack scalability, while top-down models lack mechanistic detail. A promising direction may lie in hybrid approaches that blend these strategies, enabling cross-validation of predictions and mechanisms. For instance, Miller et al. (2024) showed that specific mechanistic models (in their case, patch colonization–extinction dynamics) can be embedded within disordered systems models to explore how pairwise processes scale up in complex, realistic communities.

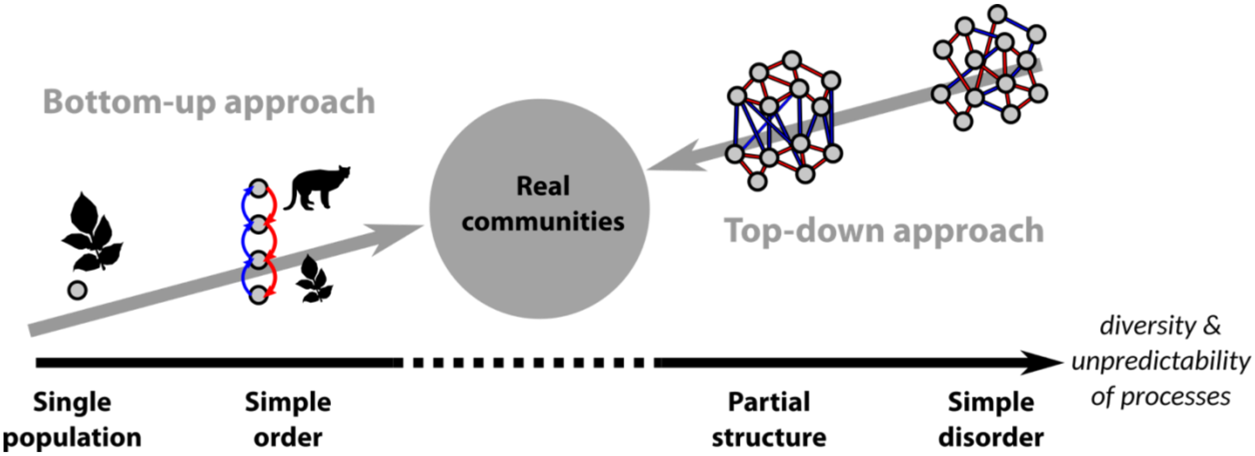

Figure 1: Two “strategic” approaches to modeling in community ecology. The more conventional “bottom-up” approach starts from the left and moves to the right. This is based on starting from single species population models, which can be arbitrarily complex to begin with, and adding species interactions in a structured way. For example, pairwise interactions may be combined to build up small sets of species as “modules” (Holt 1997). In principle, this could eventually lead to the analysis of realistically complex communities (center) that better represent those we find in nature, but progress seems to be exponentially harder as diversity increases. Alternatively, it is possible to think about communities as much more diffuse in nature. This approach is usually embodied by using random matrices to model species interactions (e.g., Random Lotka-Volterra Matrix or “RLVM” models), allowing the application of powerful mathematical methods. By progressively adding structural constraints on these models (e.g. distinct submatrices within the RLVM), they are converted into “partially structured models” that may approach realistic complex communities from another direction (right to left).

From Process to Pattern

Both the top-down and bottom-up modeling approaches can be used to explore how a limited set of processes jointly shape metacommunity patterns. One important application of either type of metacommunity model is to identify patterns associated with specific sets of mechanisms. By comparing predicted patterns linked to specific mechanisms against empirical patterns, it may be possible to diagnose the mechanisms shaping the metacommunity.

The bottom-up approach attempts to link processes to patterns in a relatively direct way. Using bottom-up models to link processes to patterns is typically most successful when few species are involved and when systems are relatively simple. Models of such systems frequently produce specific predictions that match empirical patterns both qualitatively and, in some cases, quantitatively (e.g., Vandermeer 1969; Friedman et al. 2017; Saavedra et al. 2017).

Aside from the challenge of building and analyzing models with many species, increasing complexity also makes it hard to identify any simple mapping between mechanisms and diagnostic patterns. This is an issue of convergence between models: different process-based models, even those grounded in distinct mechanisms, can produce similar or indistinguishable patterns. For example, McGill (2010) demonstrated that six mechanistically distinct models could generate similar predictions for five widely used community patterns. Consequently, none of the patterns were uniquely diagnostic of any specific model.

Given these challenges, one possible way to proceed is a "brute force" approach: associating modeled mechanisms with a larger suite of metrics or patterns (an illustrative list is shown in Table 1) to find unique combinations of predictions. This approach makes it possible to extend the one-to-one mapping between processes and patterns further into complex settings. Holling and Allen (2002) and Yanco et al. (2020) argue that evaluating models in this way can help eliminate less plausible candidates. A more targeted variation on this approach would identify a smaller subset of particularly informative or complementary metrics, which might enhance diagnostic power and increase efficiency by ruling out more mechanisms with less data.

Ultimately, the brute force approach functions as a model selection procedure: models that fail to explain the full pattern set are rejected. However, this approach does not necessarily confirm the remaining models, since new and untested alternatives may perform as well or better. For example, Ovaskainen et al. (2019), Thompson et al. (2020), and Guzman et al. (2021) applied sets of idealized and alternative models to test against a panel of pattern metrics. They found that while brute force filtering could efficiently rule out some models, it was often inconclusive in narrowing down to a single best-fit model. Including more or better metrics may help, but data will often become limiting in empirical systems.

A complementary alternative is inspired by disordered systems (top-down) modeling. Instead of seeking a one-to-one correspondence between detailed process-based models and observed patterns by enlarging the set of patterns to keep pace with the larger and larger parameter space of models, this approach seeks specific features of models that remain uniquely associated with specific patterns even as the system of interest becomes very complex. For example, Barbier et al. (2021) used data from grassland plots to predict relative yield distributions based on mean and variance in interspecific interaction strengths, linking summary statistics of interactions to a specific pattern without attempting to parameterize specific pairwise interactions. Similarly, in another study, Barbier et al. (2023) found that pairwise correlations in species abundances could reflect interaction variances and carrying capacity heterogeneity, even when actual interaction coefficients were poorly known.

Table 1: A representative set of possible metrics and patterns that can be derived from the site-by-species data matrix. Illustrative examples adapted and extended from Guzman et al. (2022). These represent a limited selection of a very large number of statistics and derived patterns that ecologists have explored and are meant to assist in our narrative of linking patterns to processes.

Metric or Pattern	Examples
Simple Descriptive Ecological Statistics	
Summaries	Means or variances of abundance, biomass, relative and absolute density, functional traits.
Distributions	Of abundances, incidences, functional traits, biomass, spatial or temporal occupancy, species-abundance distributions.
Diversity	Alpha and gamma, Hill numbers, functional and phylogenetic diversity.
Simple Derived Patterns	
Turnover	Spatial or temporal beta diversity; distance-decay, environmental-decay (with ancillary data).
Scaling relationships	Species-area, Taylor power, rarefaction or sampling curves.
Network structure	Nestedness, co-occurrence, centrality.
Second Order Derived Patterns	
Species network structure that varies as a function of environmental variation.	Metacommunity model with landscape fragmentation; time-varying trophic interactions.
Model Derived Outputs	
Variation partitioning only in space and environment	Variation partitioning of space and environment. Latent variables for each and their interactions.

Variation partitioning in time and space	Same as above but also with time.
Results from distribution models incorporating species covariances	Variation partitioning of space, environment, time, and species co-distributions; e.g., from JSDMs and similar approaches.

The key to this approach is to identify high-level parameters or parameter combinations that characterize specific ecological mechanisms and relate them to specific, diagnostic patterns. In contrast to brute force approaches, this top-down strategy targets robust, emergent features of community structure as signatures of underlying processes. However, by design, this approach cannot resolve low-level mechanistic details of metacommunity dynamics.

While this dichotomy between bottom-up and top-down approaches is simplified, it highlights contrasting philosophies in ecological modeling, and how both can be applied to link process to pattern. Hybrid strategies that integrate these perspectives, particularly through partially structured models, may offer another path forward. Appendix 1 further explores some nuances in combining these frameworks.

In sum, theory that starts with assumptions about basic processes (*sensu* Vellend 2010, 2015) and mechanisms (e.g., as reflected in assumptions about parameter distributions) does not, by itself, provide adequately resolved predictions to infer the causes that explain the patterns we see in metacommunities. The bottom-up and top-down modeling strategies suggest two distinct approaches to coping with this issue: a brute force approach, where larger sets of patterns are used to maintain the one-to-one mapping, and a disordered systems approach where one seeks higher-level mappings between summary statistics or other emergent features of models and data. Both of these approaches are inherently limited by data availability or by the ability to identify diagnostic patterns.

From Pattern to Process

Is there an alternative approach that avoids these limitations? One possibility is to use data-driven modeling methods that are less tightly linked to mechanistic assumptions. Such models can more directly decompose the different sources of variation expected to underlie community and metacommunity patterns. Statistical models aligned with Vellend's framework (2010, 2015), which decomposes community processes into selection (both biotic and abiotic), dispersal, and stochasticity (omitting speciation for now), are especially attractive when they include all of these components. Ovaskainen et al. (2017) proposed an illustrative method for doing this using Joint Species Distribution Models (JSDMs), which decompose observed variation in species distributions into components attributable to environment, space, species co-distribution (potentially indicative of interactions), and stochasticity.

Following the logic of traditional species distribution models, "classic" JSDMs apply a sequential partitioning of variation: first accounting for environment, then for additional spatial effects, and finally for residual co-distribution among species. If environmental predictors are comprehensive, any remaining co-distribution may reflect biotic interactions. Figure 2a illustrates this sequential variation partitioning.

However, JSDMs can also be used in a non-sequential framework, in which each component (environment, space, co-distribution) is estimated simultaneously. In this case, the estimated contributions correspond to marginal effects, reflecting the variation explained by each component while accounting for the presence of the others. This leads to a more complex pattern of shared and unique contributions (Figure 2b), where the interpretation of overlaps becomes ambiguous. For instance, two species might be mutually exclusive along an environmental gradient because of direct environmental filtering, competitive exclusion, or both. Without further assumptions, these effects cannot be disentangled (Dormann et al. 2018, Blanchet et al. 2018, Poggiatto et al. 2020), and they should therefore not be interpreted as interactions among predictors, but rather as overlapping marginal contributions arising from correlated or confounded sources of variation.

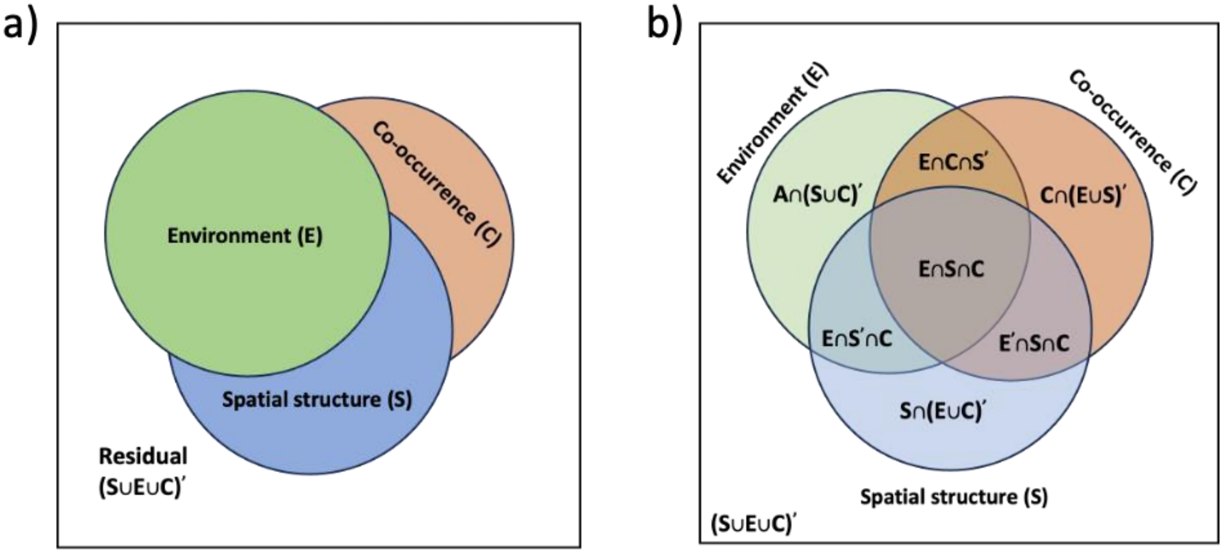

Figure 2: Variation partitioning alternatives for JSDM represented by Venn diagrams. The total variation in community composition in the metacommunity is encompassed by the outer square and is equal to 1. The white part of the figure is the unexplained or residual variation that is not accounted for by any of the predictors in the model. a) Accounting for predictors in a sequential order: First the measured environmental predictors are used and their contribution to the community variation is quantified as represented by the green area. Then spatial predictors are used to further account for community variation and their marginal contribution is shown in blue. Finally, the non-random latent correlations among the species are described, and their marginal contribution is quantified by the orange area. b) A non-sequential variation partitioning approach. Here, unique contributions from environmental, spatial, and codistribution components are represented (i.e., fractions without overlaps). Additionally there are a number of components that account for community variation that cannot be uniquely described (fractions representing predictor intersections). An interesting possibility is to compare the results of JSDM or other related methods across different metacommunities (e.g. Khattar and Peres-Neto 2024, Peak et al. 2024)

For any empirical method linking pattern to process, it is important to address uncertainty in the sampling process. Although not always implemented, several methods can help resolve this issue. These include hierarchical modeling approaches that group observations by spatial or temporal replicates to obtain sample-level statistics (Pinheiro & Bates 2000) and state-space modeling approaches that build hypothesized observation processes into the fitting algorithm (Knape and Valpine 2012) or yield "clean" datasets from which observation error has been minimized (Clark et al. 2022). Furthermore, computational limitations with very large datasets have been an important constraint on such approaches in the past, though this is increasingly being overcome (e.g., Ovaskainen et al. 2025; see Clark et al. 2018 for computationally faster alternatives). There are ongoing efforts to develop workflows that address the novel opportunities in the use of JSDMs to incorporate these developments (see Hartig et al. 2024)

A key direction for future progress may involve refining process models to better align with the data structures used in pattern-based inference. For example, Leibold et al. (2022) proposed analyzing the internal structure of metacommunities by decomposing JSDM components across species and sites. This decomposition reveals how different species respond to environmental and spatial variation and how their distributions co-vary. Such an approach is promising because it moves beyond global summaries to examine heterogeneity in processes across species and locations, thereby allowing more direct links between ecological mechanisms and observed patterns. By identifying which species and sites drive particular components of variation, these methods can generate testable hypotheses about underlying processes and help bridge the gap between pattern-based inference and process-based ecological understanding. Unlike traditional pattern metrics, which aggregate across species, this approach allows process inference to operate at the species- and/or patch- level. In doing so, it opens the possibility of aligning species-specific effects with bottom-up models that operate on subsets of interacting taxa. This could eventually allow for a robust integration of bottom-up and top-down perspectives, as outlined in Figure 3.

Figure 3 illustrates potential inference pipelines that connect process-based models and data-based metrics. The top half of the figure emphasizes aggregated descriptors (e.g., means, variances), while the bottom half highlights models with more detailed parameterization. Arrows show how models and metrics can reject or support hypotheses, and where gaps in inference remain, particularly the persistent entanglement of processes such as dispersal, environmental filtering, and species interactions.

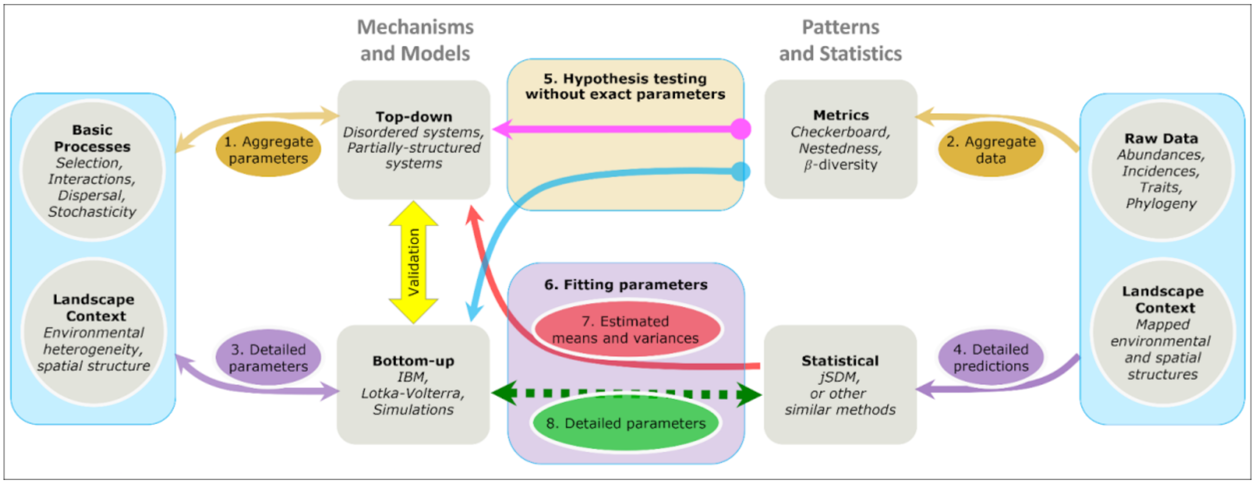

Figure 3: Pipelines between 'Process' (left) and 'Pattern' (right). We distinguish between approaches that aggregate parameters and/or data (upper pipeline in gold) and those that focus on maintaining detailed parameters and predictors (lower pipeline in violet). We define process-based approaches (starting from the left) as those that combine basic processes (i.e. density-independent selection, density dependent selection, dispersal, stochasticity) to make predictions about resulting patterns within landscapes that consider levels of heterogeneity and spatial structure (left side of the figure). The 'top-down approach' (1) uses 'disordered' or 'partially structured' models by aggregating parameters (typically, into mean, variances and covariances of these parameters) to predict robust patterns, but sacrifices predicting 'fragile' detailed components of those patterns. Alternatively, there is a 'bottom-up' approach (2) that specifies parameters, typically in models with relatively few species, and uses detailed model specifications (e.g., using individually based models (IBMS), Lotka-Volterra models with specified parameters, or other simulation models) to make detailed predictions. Such models can also predict some of the same general patterns as the top-down models but aim to be able to make more detailed predictions that may produce detailed fit to data. There is a parallel structure that derives patterns from data (typically the site-by-species matrix, along with ancillary data on species traits/interrelations and/or explicit landscape features). One approach derives aggregated pattern descriptors (gold arrow 2). These include a wide array of 'metrics' such as species area relations, diversity and related methods, etc. A more recent effort has been directed at identifying parameter values from data. These metrics provide means of rejecting particular hypotheses (5. and 6., magenta and blue lines) but do not normally parameterize the processes. An alternate approaches uses sophisticated statistical modeling to identify detailed features of mechanisms. This may involve either the inferences of statistical features (red arrow, 7) via e.g. JSDMs to infer metacommunity mechanisms (e.g. Ovaskainen et al. 2017, Leibold et al. 2022) or an attempt to derice actual parameter values (green dashed arrow, 8). These approaches, while desirable, have not yet been able to resolve the entanglement of the five processes to clearly parameterize mechanistic models. In the absence of such detailed inference of models. An intriguing point is to explore how bottom-up and top-down methods might be cross-validated (yellow arrow).

In summary, the mapping from patterns to processes remains complex and often ambiguous. However, methodological advances, particularly JSDMs and species-level decomposition, offer a promising route forward. Further refinement of both models and inference tools may help to close the gap between observed community patterns and the underlying processes that generate them.

Conclusions and Prospectus

We have outlined the conceptual and methodological challenges involved in linking processes and patterns in ecological communities and metacommunities. These issues trace back to the origins of community ecology, yet recent decades have seen tremendous progress. The development of metacommunity theory has notably reshaped how we interpret spatial biodiversity patterns (Leibold et al. 2004, Holyoak et al. 2005, Leibold and Chase 2018). On the analytical side, species distribution models have evolved from environmentally focused predictors to those incorporating spatial effects and species co-distributions, thereby often increasing predictive power from roughly 20–30% to 70–80% in some cases (Leibold and Peres-Neto unpublished results).These gains arise from incorporating a much broader set of predictors, including complex interactions among species co-distributions, environmental conditions, and robust spatial predictors. Although such models can become difficult to interpret mechanistically, their improved predictive performance offers hope that integrating multiple sources of variation can help reveal the dominant processes structuring communities and guide the development of more interpretable, process-based frameworks.

Despite these advances, substantial challenges remain. Chief among them is the uncertainty of current inference approaches. This problem is especially clear in JSDMs, for example, where the effects of density-independent environmental filtering and density-dependent biotic interactions are often conflated. Similar ambiguity arises in multi-metric brute force approaches, as illustrated by McGill (2010), and in simulation studies (Chave et al. 2002, Ovaskainen et al. 2019; Thompson et al. 2020; Guzman et al. 2021) that failed to uniquely support any single mechanistic model.

It is important to acknowledge that many existing methods were not designed to infer mechanisms with high specificity. Often, the goal has been to reject null models or generate predictive models without mechanistic interpretation. While such approaches are valuable, particularly in applied contexts, the need for mechanistically grounded prediction grows as we face novel environmental conditions.

Encouragingly, the field is evolving. Developments in data collection, modeling, and computation suggest a promising trajectory rather than a fixed endpoint. We anticipate major progress in the coming years, driven by:

- **Improved data streams** that enhance sampling coverage, temporal resolution, and accuracy, including being able to address uncertainty in sampling.
- **Greater computational capacity** to analyze complex datasets and run more sophisticated models, including those incorporating machine learning or artificial intelligence.
- **Advances in process modeling**, both in modular bottom-up approaches (e.g., for eco-evolutionary dynamics) and in statistical-mechanics-based top-down approaches, particularly through partially structured models.

- **Increased emphasis on validation and prediction**, to better link model outputs with empirical data and strengthen the integration of metrics and models.

To accelerate progress, we suggest several promising directions:

1. Integrate temporal dynamics and multi-scale data: Most metacommunity data remain spatial, but temporal metrics offer powerful insights (Holyoak et al. 2020, Record et al. 2021, Gálvez et al. 2024). For example, Jabot et al. (2020) and Guzman et al. (2022) showed how accounting for temporal effects improves the resolution of metacommunity structure. Temporal data also facilitate causal inference via concepts like Granger causality, which can help infer directionality in species responses; something not possible from spatial data alone. Jackson et al. (2025) propose a novel and promising analytical method that exploits this aspect of metacommunity analysis called "maximum caliber".
2. **Design and incorporate experimental manipulations**: Experiments can validate observational inferences and provide stronger process-based insights (Werner 1998; Grace 2024). Effective experimental designs, such as manipulating species presence/absence across environments, could directly test candidate mechanisms identified in models. Excitingly, advances in sequencing and high-throughput culture methods have fostered an explosive growth in experimental microbial ecology and the unique tractability of microbial (meta)communities makes them a promising platform for such experimental manipulations. Approaches that combine experimental, observational, and theoretical approaches to address key conceptual issues are particularly compelling (e.g. Abrego et al. 2025, Hu et al. 2025).
3. **Bridge top-down and bottom-up approaches**: Rather than treating these as oppositional approaches, hybrid strategies can combine mechanistic insights from modules with the robustness of statistical aggregation. Partially structured models are one way to integrate these perspectives. For instance, Barbier et al. (2018) showed how modest structure imposed on random matrices could shift predictions significantly, combining both realism and analytical tractability.

Community ecology, and particularly metacommunity theory, is undergoing methodological and conceptual transformations. Understanding the links between processes and patterns remains a central goal. Doing so not only enhances ecological theory but also improves prediction under environmental change. Purely predictive models will struggle as conditions diverge from historical baselines. Mechanistic understanding, while harder to achieve, provides a necessary complement to maintain predictive power. By assessing current approaches and their limitations and highlighting promising future directions, we aim to support and accelerate progress in this critical area. Ultimately, bridging the gap between pattern and process is not only a heuristic ambition, but also an essential step toward predictive and causal ecology. By leveraging the complementary strengths of diverse approaches and fostering integration across empirical, theoretical, and computational fronts, we can move closer to a unified framework for understanding the dynamics of biodiversity.

Speculation and Alternative Viewpoints

Process-to-pattern and pattern-to-process share the same overall goals but reflect different mindsets and practices and may thus end up in different places. Starting with processes can yield patterns that are not necessarily observable: for instance, much theoretical work on stability, including May's, is only usable qualitatively because its main quantitative predictions are eigenvalues that cannot be measured. On the other hand, starting with patterns can focus our attention on salient empirical regularities, such as Species Abundance Distributions, that are poorly diagnostic of underlying processes (many different mechanisms yielding very similar distributions). In addition, pattern-oriented modelling tends to make assumptions about processes that are more limiting, such as assuming a simple additive model of contributions from various processes, enabling variance partitioning, whereas a process-based dynamical model can display a much more complex entanglement of these causes.

Matthieu Barbier

I think the way that we are most likely to move forward to address this long-standing issue in ecology of connecting process to pattern as well as pattern to process is to have theoretical modelers and empirical researchers design experiments together from the beginning, making sure that the relevant parameters will be able to be measured and that actual real experimental results can test predictions from models. Too often we are collaborating after the fact, which hobbles both sides. But this approach is hard because it takes more time up front and means that compromises need to be made.

Leonora Bittleston

While the issue of pattern-to-process and process-to-pattern can sound technically overwhelming, this topic is of grave importance for ecological applications and science communication. As described above, the plethora of modeling approaches available to ecologists follow different work streams. Yet, despite their differences, terms like "importance", "significance", "driver", and "effect" are commonly used across virtually all of them. Such words risk getting conflated with interpretative words like "cause" or "mechanism", and it is here that applications and decision-making are the most at risk. In identifying where our work stands across the pattern-to-process / process-to-pattern pipeline as a first step, we can develop cautionary guard-rails that help to clearly communicate our research with stakeholders, governments, the general public, and even fellow ecologists.

Veronica F. Frans

While analytically intractable, multispecies process-based models can be analyzed numerically to study links between pattern and process in large communities. This is particularly useful when analytic results do not generalize beyond two species. One example is the use of species trait distributions to infer the relative roles of biotic and abiotic selection. Computational models show that competition can cause trait-based clustering in multispecies communities (Scheffer et al. 2006, D'Andrea et al. 2019,), a pattern traditionally attributed to environmental filtering (Gotzenberger et al. 2012). Although such models are unlikely to produce precise predictions linking species abundances to trait values, coarse-grained patterns such as competition-induced trait clustering may be expected in hyperdiverse systems subject to intense competition for limited resources, such as tropical forests and microbial communities (D'Andrea et al. 2020, Martinez=Blancas et al. 2022).

Rafael D'Andrea

One issue that deserves emphasis is that the challenge often lies not in the failure of an inferential tool, but in how its results are interpreted. Variation partitioning illustrates a broader class of pattern-based, bottom-up inference methods that decompose community variation into components associated with environment, space, and residual structure. Its strength lies not in uniquely identifying mechanisms, but in providing a structured summary of complex metacommunity data and clarifying what remains unexplained. This is particularly useful for detecting environmental selection, accounting for spatial autocorrelation, and identifying structured residual variation that may reflect missing predictors, historical contingencies, or unmodeled interactions. Difficulties arise when these broad components are overinterpreted mechanistically. For example, when spatial structure is taken as direct evidence of dispersal or neutral dynamics. This limitation is not unique to variation partitioning but applies broadly to pattern-based statistical decompositions. Rather than viewing such methods as inadequate, they are best seen as important first steps in a broader inferential pipeline, helping organize signals in the data, constrain plausible explanations, and guide the development of more explicit mechanistic models.

Pedro R. Peres-Neto

Acknowledgements

This study was conceived and executed by a working group of participants of the National Science Foundation-funded workshop, “Merging Statistical Theory and Analyses at the Interface of Microbial and Macrobial Ecology” (Award Number: NSF2224331, Montreal, Canada, 2022). VFF was supported by the NSF Postdoctoral Research Fellowship in Biology (NSF Grant no. 2410350) and the Stanford Science Fellows program (Stanford University), MAL was also supported by NSF AWD08828.

Literature Cited:

Abrego, N., Saine, S., Penttilä, R., Furneaux, B., Hytönen, T., Miettinen, O., Monkhouse, N., Mäkipää, R., Pennanen, J., Zakharov, E. V., & Ovaskainen, O. (2025). The role of stochasticity in fungal community assembly: Explaining apparent stochasticity with field experiments. *Proceedings of the Royal Society B: Biological Sciences*, *292*(2040), 20242416. https://doi.org/10.1098/rspb.2024.2416

Ahmadian, Y., Fumarola, F., & Miller, K. D. (2015). Properties of networks with partially structured and partially random connectivity. *Physical Review E*, *91*(1), 012820. https://doi.org/10.1103/PhysRevE.91.012820

Akjouj, I., Barbier, M., Clenet, M., Hachem, W., Maïda, M., Massol, F., Najim, J., & Tran, V. C. (2024). Complex systems in ecology: A guided tour with large Lotka–Volterra models and random matrices. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *480*(2285), 20230284. https://doi.org/10.1098/rspa.2023.0284

Amarasekare, P. (2003). Competitive coexistence in spatially structured environments: A synthesis. *Ecology Letters*, *6*(12), 1109–1122. https://doi.org/10.1046/j.1461-0248.2003.00530.x

Amarasekare, P., & Nisbet, R. M. (2001). Spatial Heterogeneity, Source-Sink Dynamics, and the Local Coexistence of Competing Species. *The American Naturalist*, *158*(6), 572–584. https://doi.org/10.1086/323586

Barbier, M., Arnoldi, J.-F., Bunin, G., & Loreau, M. (2018). Generic assembly patterns in complex ecological communities. *Proceedings of the National Academy of Sciences*, *115*(9), 2156–2161. https://doi.org/10.1073/pnas.1710352115

Barbier, M., Bunin, G., & Leibold, M. A. (2023). *Getting More by Asking for Less: Linking Species Interactions to Species Co-Distributions in Metacommunities*. Ecology. https://doi.org/10.1101/2023.06.04.543606

Barbier, M., De Mazancourt, C., Loreau, M., & Bunin, G. (2021). Fingerprints of High-Dimensional Coexistence in Complex Ecosystems. *Physical Review X*, *11*(1), 011009. https://doi.org/10.1103/PhysRevX.11.011009

Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, *23*(7), 1050–1063. https://doi.org/10.1111/ele.13525

Borcard, D., Legendre, P., & Drapeau, P. (1992). Partialling out the Spatial Component of Ecological Variation. *Ecology*, *73*(3), 1045–1055. https://doi.org/10.2307/1940179

Borregaard, M. K., Rahbek, C., Fjeldså, J., Parra, J. L., Whittaker, R. J., & Graham, C. H. (2014). Node-based analysis of species distributions. *Methods in Ecology and Evolution*, *5*(11), 1225–1235. https://doi.org/10.1111/2041-210X.12283

Borthagaray, A. I., Arim, M., & Marquet, P. A. (2014). Inferring species roles in metacommunity structure from species co-occurrence networks. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1792), 20141425. https://doi.org/10.1098/rspb.2014.1425

Carugno, G., Neri, I., & Vivo, P. (2022). Instabilities of complex fluids with partially structured and partially random interactions. *Physical Biology*, *19*(5), 056001. https://doi.org/10.1088/1478-3975/ac55f9

Chang, C.-Y., Bajić, D., Vila, J. C. C., Estrela, S., & Sanchez, A. (2023). Emergent coexistence in multispecies microbial communities. *Science*, *381*(6655), 343–348. https://doi.org/10.1126/science.adg0727

Chave, J., Muller-Landau, H. C., & Levin, S. A. (2002). Comparing Classical Community Models: Theoretical Consequences for Patterns of Diversity. *The American Naturalist*, *159*(1), 1–23. https://doi.org/10.1086/324112

Choler, P., Michalet, R., & Callaway, R. M. (2001). FACILITATION AND COMPETITION ON GRADIENTS IN ALPINE PLANT COMMUNITIES. *Ecology*, *82*(12), 3295–3308. https://doi.org/10.1890/0012-9658(2001)082%255B3295:FACOGI%255D2.0.CO;2

Clark, A. T., Lehman, C., & Tilman, D. (2018). Identifying mechanisms that structure ecological communities by snapping model parameters to empirically observed tradeoffs. *Ecology Letters*, *21*(4), 494–505. https://doi.org/10.1111/ele.12910

Clark, A. T., Mühlbauer, L. K., Hillebrand, H., & Karakoç, C. (2022). Measuring stability in ecological systems without static equilibria. *Ecosphere*, *13*(12), e4328. https://doi.org/10.1002/ecs2.4328

Connell, J. H. (1983). On the Prevalence and Relative Importance of Interspecific Competition: Evidence from Field Experiments. *The American Naturalist*, *122*(5), 661–696. https://doi.org/10.1086/284165

Connor, E. F., & Simberloff, D. (1979). The Assembly of Species Communities: Chance or Competition? *Ecology*, *60*(6), 1132–1140. https://doi.org/10.2307/1936961

Corrado, L. (2005). Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models. *Journal of the American Statistical Association*, *100*(470), 710–711. https://doi.org/10.1198/jasa.2005.s25

Cottenie, K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, *8*(11), 1175–1182. https://doi.org/10.1111/j.1461-0248.2005.00820.x

Crombie, A. C. (n.d.). On competition between different species of graniivorous insects. *Proceedings of the Royal Society B: Biological Sciences*.

Crombie, A. C. (1946). Further experiments on insect competition. *Proceedings of the Royal Society of London. Series B - Biological Sciences*, *133*(870), 76–109. https://doi.org/10.1098/rspb.1946.0004

Cui, W., Marsland, R., & Mehta, P. (2024). *Les Houches Lectures on Community Ecology: From Niche Theory to Statistical Mechanics* (No. arXiv:2403.05497). arXiv. https://doi.org/10.48550/arXiv.2403.05497

D’Andrea, R., Guittar, J., O’Dwyer, J. P., Figueroa, H., Wright, S. J., Condit, R., & Ostling, A. (2020). Counting niches: Abundance-by-trait patterns reveal niche partitioning in a Neotropical forest. *Ecology*, *101*(6), e03019. https://doi.org/10.1002/ecy.3019

D’Andrea, R., Riolo, M., & Ostling, A. M. (2019). Generalizing clusters of similar species as a signature of coexistence under competition. *PLOS Computational Biology*, *15*(1), e1006688. https://doi.org/10.1371/journal.pcbi.1006688

Darwin, C. (1859). *On the Origin of Species*. John Murray.

Dormann, C. F., Bobrowski, M., Dehling, D. M., Harris, D. J., Hartig, F., Lischke, H., Moretti, M. D., Pagel, J., Pinkert, S., Schleuning, M., Schmidt, S. I., Sheppard, C. S., Steinbauer, M. J., Zeuss, D., & Kraan, C. (2018). Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. *Global Ecology and Biogeography*, *27*(9), 1004–1016. https://doi.org/10.1111/geb.12759

Friedman, J., Higgins, L. M., & Gore, J. (2017). Community structure follows simple assembly rules in microbial microcosms. *Nature Ecology & Evolution*, *1*(5), 0109. https://doi.org/10.1038/s41559-017-0109

Gálvez, Á., Peres-Neto, P. R., Castillo-Escrivà, A., Bonilla, F., Camacho, A., García-Roger, E. M., Iepure, S., Miralles, J., Monrós, J. S., Olmo, C., Picazo, A., Rojo, C., Rueda, J., Sasa, M., Segura, M., Armengol, X., & Mesquita-Joanes, F. (2024). Spatial versus spatio-temporal approaches for studying metacommunities: A multi-taxon analysis in Mediterranean and tropical temporary ponds. *Proceedings of the Royal Society B: Biological Sciences*, *291*(2020), 20232768. https://doi.org/10.1098/rspb.2023.2768

Gause, G. F. (1932). *Experimental Studies on the Struggle for Existence In Mixed Population of Two Species of Yeast*. *9*, 389–402.

Gause, G. F. (1934). *The Struggle for Existence*. Williams and Wilkins.

Germain, R. M., Hart, S. P., Turcotte, M. M., Otto, S. P., Sakarchi, J., Rolland, J., Usui, T., Angert, A. L., Schluter, D., Bassar, R. D., Waters, M. T., Henao-Diaz, F., & Siepielski, A. M. (2021). On the Origin of Coexisting Species. *Trends in Ecology & Evolution*, *36*(4), 284–293. https://doi.org/10.1016/j.tree.2020.11.006

Goh, B. S. (1979). Stability in models of mutualism. *The American Naturalist*, *113*, 261–275.

Gotelli, N. J., & McCabe, D. J. (2002). SPECIES CO-OCCURRENCE: A META-ANALYSIS OF J. M. DIAMOND'S ASSEMBLY RULES MODEL. *Ecology*, *83*(8), 2091–2096. https://doi.org/10.1890/0012-9658(2002)083%255B2091:SCOAMA%255D2.0.CO;2

Götzenberger, L., De Bello, F., Bråthen, K. A., Davison, J., Dubuis, A., Guisan, A., Lepš, J., Lindborg, R., Moora, M., Pärtel, M., Pellissier, L., Pottier, J., Vittoz, P., Zobel, K., & Zobel, M. (2012). Ecological assembly rules in plant communities—Approaches, patterns and prospects. *Biological Reviews*, *87*(1), 111–127. https://doi.org/10.1111/j.1469-185X.2011.00187.x

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

Grace, J. B., Harrison, S., & Damschen, E. I. (2011). Local richness along gradients in the Siskiyou herb flora: R. H. Whittaker revisited. *Ecology*, *92*(1), 108–120. https://doi.org/10.1890/09-2137.1

Gravel, D., Massol, F., & Leibold, M. A. (2016). Stability and complexity in model meta-ecosystems. *Nature Communications*, *7*(1), 12457. https://doi.org/10.1038/ncomms12457

Gurevitch, J., Morrow, L. L., Wallace, A., & Walsh, J. S. (1992). A Meta-Analysis of Competition in Field Experiments. *The American Naturalist*, *140*(4), 539–572. https://doi.org/10.1086/285428

Guzman, L. M., Thompson, P. L., Viana, D. S., Vanschoenwinkel, B., Horváth, Z., Ptacnik, R., Jeliazkov, A., Gascón, S., Lemmens, P., Anton-Pardo, M., Langenheder, S., De Meester, L., & Chase, J. M. (2022). Accounting for temporal change in multiple biodiversity patterns improves the inference of metacommunity processes. *Ecology*, *103*(6), e3683. https://doi.org/10.1002/ecy.3683

Hart, S. P., Freckleton, R. P., & Levine, J. M. (2018). How to quantify competitive ability. *Journal of Ecology*, *106*(5), 1902–1909. https://doi.org/10.1111/1365-2745.12954

Harte, J., & Newman, E. A. (2014). Maximum information entropy: A foundation for ecological theory. *Trends in Ecology & Evolution*, *29*(7), 384–389. https://doi.org/10.1016/j.tree.2014.04.009

Hartig, F., Abrego, N., Bush, A., Chase, J. M., Guillera-Arroita, G., Leibold, M. A., Ovaskainen, O., Pellissier, L., Pichler, M., Poggiato, G., Pollock, L., Si-Moussi, S., Thuiller, W., Viana, D. S., Warton, D. I., Zurell, D., & Yu, D. W. (2024). Novel community data in ecology-properties and prospects. *Trends in Ecology & Evolution*, *39*(3), 280–293. https://doi.org/10.1016/j.tree.2023.09.017

Holling, C. S., & Allen, C. R. (2002a). Adaptive Inference for Distinguishing Credible from Incredible Patterns in Nature. *Ecosystems*, *5*(4), 319–328. https://doi.org/10.1007/s10021-001-0076-2

Holling, C. S., & Allen, C. R. (2002b). Adaptive Inference for Distinguishing Credible from Incredible Patterns in Nature. *Ecosystems*, *5*(4), 319–328. https://doi.org/10.1007/s10021-001-0076-2

Holt, R. D., Grover, J., & Tilman, D. (1994). Simple Rules for Interspecific Dominance in Systems with Exploitative and Apparent Competition. *The American Naturalist*, *144*(5), 741–771. https://doi.org/10.1086/285705

Holt, R. D., & Hochberg, M. E. (2001). Indirect interactions, community modules and biological control: A theoretical perspective. In E. Wajnberg, J. K. Scott, & P. C. Quimby (Eds.), *Evaluating indirect ecological effects of biological control. Key papers from the symposium "Indirect ecological effects in biological control", Montpellier, France, 17-20 October 1999* (1st ed., pp. 13–37). CABI Publishing. https://doi.org/10.1079/9780851994536.0013

Holyoak, M., Caspi, T., & Redosh, L. W. (2020). Integrating Disturbance, Seasonality, Multi-Year Temporal Dynamics, and Dormancy Into the Dynamics and Conservation of Metacommunities. *Frontiers in Ecology and Evolution*, *8*, 571130. https://doi.org/10.3389/fevo.2020.571130

Holyoak, M., Leibold, M. A., & Holt, R. D. (Eds.). (2005). *Metacommunities: Spatial dynamics and ecological communities*. University of Chicago Press.

Hu, J., Barbier, M., Bunin, G., & Gore, J. (2025). Collective dynamical regimes predict invasion success and impacts in microbial communities. *Nature Ecology & Evolution*, *9*(3), 406–416. https://doi.org/10.1038/s41559-024-02618-y

Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton University Press.

Huffaker, C. B. (1958). Experimental studies on predation: Dispersion factors and predator-prey oscillations. *Hilgardia*, *27*(14), 343–383. https://doi.org/10.3733/hilg.v27n14p343

Jabot, F., Laroche, F., Massol, F., Arthaud, F., Crabot, J., Dubart, M., Blanchet, S., Munoz, F., David, P., & Datry, T. (2020). Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition. *Ecology Letters*, *23*(9), 1330–1339. https://doi.org/10.1111/ele.13523

Jackson, Z., Leibold, M. A., Holt, R. D., & Xue, B. (2026). Modeling and inferring metacommunity dynamics with Maximum Caliber. *Proceedings of the National Academy of Sciences*, *123*(1), e2520867123. https://doi.org/10.1073/pnas.2520867123

Kéfi, S., Miele, V., Wieters, E. A., Navarrete, S. A., & Berlow, E. L. (2016). How Structured Is the Entangled Bank? The Surprisingly Simple Organization of Multiplex Ecological Networks Leads to Increased Persistence and Resilience. *PLOS Biology*, *14*(8), e1002527. https://doi.org/10.1371/journal.pbio.1002527

Knape, J., & De Valpine, P. (2012). Fitting complex population models by combining particle filters with Markov chain Monte Carlo. *Ecology, 93*(2), 256–263. https://doi.org/10.1890/11-0797.1

Kneitel, J. M., & Chase, J. M. (2004). Trade-offs in community ecology: Linking spatial scales and species coexistence. *Ecology Letters*, *7*(1), 69–80. https://doi.org/10.1046/j.1461-0248.2003.00551.x

Leibold, M. A. (1996). A Graphical Model of Keystone Predators in Food Webs: Trophic Regulation of Abundance, Incidence, and Diversity Patterns in Communities. *The American Naturalist*, *147*(5), 784–812. https://doi.org/10.1086/285879

Leibold, M. A., & Chase, J. M. (2018). *Metacommunity Ecology, Volume 59*. Princeton University Press. https://doi.org/10.1515/9781400889068

Leibold, M. A., Govaert, L., Loeuille, N., De Meester, L., & Urban, M. C. (2022). Evolution and Community Assembly Across Spatial Scales. *Annual Review of Ecology, Evolution, and Systematics*, *53*(1), 299–326. https://doi.org/10.1146/annurev-ecolsys-102220-024934

Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., & Gonzalez, A. (2004). The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, *7*(7), 601–613. https://doi.org/10.1111/j.1461-0248.2004.00608.x

Leibold, M. A., & Mikkelson, G. M. (2002). Coherence, species turnover, and boundary clumping: Elements of meta-community structure. *Oikos*, *97*(2), 237–250. https://doi.org/10.1034/j.1600-0706.2002.970210.x

Leibold, M. A., Rudolph, F. J., Blanchet, F. G., De Meester, L., Gravel, D., Hartig, F., Peres-Neto, P., Shoemaker, L., & Chase, J. M. (2022). The internal structure of metacommunities. *Oikos*, *2022*(1), oik.08618. https://doi.org/10.1111/oik.08618

Lerch, B. A., Rudrapatna, A., Rabi, N., Wickman, J., Koffel, T., & Klausmeier, C. A. (2023). Connecting local and regional scales with stochastic metacommunity models: Competition, ecological drift, and dispersal. *Ecological Monographs*, *93*(4), e1591. https://doi.org/10.1002/ecm.1591

Lotka, A. (1925). *Elements of Physical Biology*. Williams and Wilkins.

MacArthur, R. H. (1974). *Geographical Ecology: Patterns in the Distribution of Species*.

MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton University Press.

Martínez-Blancas, A., Belaustegui, I. X., & Martorell, C. (2022). Species alliances and hidden niche dimensions drive species clustering along a hydric gradient in a semiarid grassland. *Ecology Letters*, *25*(12), 2651–2662. https://doi.org/10.1111/ele.14122

May, F., Huth, A., & Wiegand, T. (2015). Moving beyond abundance distributions: Neutral theory and spatial patterns in a tropical forest. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1802), 20141657. https://doi.org/10.1098/rspb.2014.1657

May, F., Wiegand, T., Lehmann, S., & Huth, A. (2016). Do abundance distributions and species aggregation correctly predict macroecological biodiversity patterns in tropical forests? *Global Ecology and Biogeography*, *25*(5), 575–585. https://doi.org/10.1111/geb.12438

May, R. M. (1972). Will a large complex system be stable? *Nature*, *238*, 413–414.

May, R. M. (1973). *Stability and Complexity in Model Ecosystems.* . (Vol. 6). Princeton University Press.

McGill, B. J. (2010). Towards a unification of unified theories of biodiversity. *Ecology Letters*, *13*(5), 627–642. https://doi.org/10.1111/j.1461-0248.2010.01449.x

McPeek, M. A. (2022). *Coexistencein ecology: A mechanistic perspective*. Princeton University Press.

Miller, Z. R., Clenet, M., Della Libera, K., Massol, F., & Allesina, S. (2024). Coexistence of many species under a random competition–colonization trade-off. *Proceedings of the National Academy of Sciences*, *121*(5), e2314215121. https://doi.org/10.1073/pnas.2314215121

Mühlbauer, L. K., Schulze, M., Harpole, W. S., & Clark, A. T. (2020). gauseR: Simple methods for fitting Lotka-Volterra models describing Gause's "Struggle for Existence." *Ecology and Evolution*, *10*(23), 13275–13283. https://doi.org/10.1002/ece3.6926

Ovaskainen, O., Rybicki, J., & Abrego, N. (2019). What can observational data reveal about metacommunity processes? *Ecography*, *42*(11), 1877–1886. https://doi.org/10.1111/ecog.04444

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., & Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, *20*(5), 561–576. https://doi.org/10.1111/ele.12757

Ovaskainen, O., Winter, S., Tikhonov, G., Abrego, N., Anslan, S., deWaard, J. R., deWaard, S. L., Fisher, B. L., Furneaux, B., Hardwick, B., Kerdraon, D., Pentinsaari, M., Raharinjanahary, D., Rajoelison, E. T., Ratnasingham, S., Somervuo, P., Sones, J. E., Zakharov, E. V., Hebert, P. D. N., … Dunson, D. (2025). Common to rare transfer learning

(CORAL) enables inference and prediction for a quarter million rare Malagasy arthropods. *Nature Methods*, *22*(10), 2074–2082. https://doi.org/10.1038/s41592-025-02823-y

Pascual, M. A., & Kareiva, P. (1996). Predicting the outcome of competition using experimental data: Maximum likelihood and bayesian approaches. *Ecology*, *77*, 337–249.

Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). VARIATION PARTITIONING OF SPECIES DATA MATRICES: ESTIMATION AND COMPARISON OF FRACTIONS. *Ecology*, *87*(10), 2614–2625. https://doi.org/10.1890/0012-9658(2006)87%255B2614:VPOSDM%255D2.0.CO;2

Peres-Neto, P. R., Olden, J. D., & Jackson, D. A. (2001). Environmentally constrained null models: Site suitability as occupancy criterion. *Oikos*, *93*(1), 110–120. https://doi.org/10.1034/j.1600-0706.2001.930112.x

Pinheiro, J. C., & Bates, D. M. (2004). *Mixed-effects models in S and S-PLUS* (Nachdr.). Springer.

Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J. S., & Thuiller, W. (2021). On the Interpretations of Joint Modeling in Community Ecology. *Trends in Ecology & Evolution*, *36*(5), 391–401. https://doi.org/10.1016/j.tree.2021.01.002

Ranjan, R., Koffel, T., & Klausmeier, C. A. (2024). The three-species problem: Incorporating competitive asymmetry and intransitivity in modern coexistence theory. *Ecology Letters*, *27*(4), e14426. https://doi.org/10.1111/ele.14426

Saavedra, S., Rohr, R. P., Bascompte, J., Godoy, O., Kraft, N. J. B., & Levine, J. M. (2017). A structural approach for understanding multispecies coexistence. *Ecological Monographs*, *87*(3), 470–486. https://doi.org/10.1002/ecm.1263

Sanderson, J. G., & Pimm, S. L. (2015). *Patterns in Nature: The Analysis of Species Co-occurences* (s). University of Chicago Press.

Schaffer, W. M. (1981). Ecological Abstraction: The Consequences of Reduced Dimensionality in Ecological Models. *Ecological Monographs*, *51*(4), 383–401. https://doi.org/10.2307/2937321

Scheffer, M., & Van Nes, E. H. (2006). Self-organized similarity, the evolutionary emergence of groups of similar species. *Proceedings of the National Academy of Sciences*, *103*(16), 6230–6235. https://doi.org/10.1073/pnas.0508024103

Schoener, T. W. (1983). Field Experiments on Interspecific Competition. *The American Naturalist*, *122*(2), 240–285. https://doi.org/10.1086/284133

Serván, C. A., Capitán, J. A., Miller, Z. R., & Allesina, S. (2025). Effects of Phylogeny on Coexistence in Model Communities. *The American Naturalist*, *205*(2), E34–E48. https://doi.org/10.1086/733415

Shoemaker, W. R., Sánchez, Á., & Grilli, J. (2023). *Macroecological patterns in experimental microbial communities*. https://doi.org/10.1101/2023.07.24.550281

Sih, A., Crowley, P., McPeek, M., Petranka, J., & Strohmeier, K. (1985). Predation, Competition, and Prey Communities: A Review of Field Experiments. *Annual Review of Ecology and Systematics*, *16*(1), 269–311. https://doi.org/10.1146/annurev.es.16.110185.001413

Skrondal, A. (with Rabe-Hesketh, S.). (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Chapman & Hall/CRC.

Ter Braak, C. J. F., & Prentice, I. C. (1988). A Theory of Gradient Analysis. In *Advances in Ecological Research* (Vol. 18, pp. 271–317). Elsevier. https://doi.org/10.1016/S0065-2504(08)60183-X

Thom, R., Noel, E. (2016). *To predict is not to explain: Conversations on mathematics, science, catastrophe theory, semiophysics, natural philosophy and morphogenesis*. Thobooks Press.

Thompson, P. L., Guzman, L. M., De Meester, L., Horváth, Z., Ptacnik, R., Vanschoenwinkel, B., Viana, D. S., & Chase, J. M. (2020). A process-based metacommunity framework linking local and regional scale community ecology. *Ecology Letters*, *23*(9), 1314–1329. https://doi.org/10.1111/ele.13568

Tilman, D. (1982). *Resource Competition and Community Structure*.

Vandermeer, J. H. (1969). The Competitive Structure of Communities: An Experimental Approach with Protozoa. *Ecology*, *50*(3), 362–371. https://doi.org/10.2307/1933884

Werner, E. E. (1998). Ecological experiments and a research program in community ecology. In *Experimental Ecology: Issues and Perspectives* (pp. 3–26). Oxford University Press.

Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, *30*(3), 279–338. https://doi.org/10.2307/1943563

Yanco, S. W., McDevitt, A., Trueman, C. N., Hartley, L., & Wunder, M. B. (2020). A modern method of multiple working hypotheses to improve inference in ecology. *Royal Society Open Science*, *7*(6), 200231. https://doi.org/10.1098/rsos.200231

Yodzis, P. (1981). The stability of real ecosystems. *Nature*, *289*(5799), 674–676. https://doi.org/10.1038/289674a0

Supplemental Information

Appendix 1: Glossary of terms:

Table S1. An annotated glossary of terms used in this essay. Their intended definition is to refine how we link patterns to processes in metacommunities (here, our system of study).

Term	Definition
Data	Direct observations of the focal elements in a system. Here, usually represented by a site-by-species-by-time matrix of presence/absence, abundance/biomass, or relative abundance/biomass.
Pattern	Informative or interesting regularities within a system often described as simplified representations of data.
Mechanism	Specific configurations of processes and constraints driving processes that produce particular or relevant outcomes. Mechanisms enable the stages of a process to occur by providing the elements necessary for each step in a process.
Processes	Fundamental elements of the ecological dynamics of a system. A series of steps necessary to achieve a particular pattern or behaviour in a system. These are reasonably well summarized by Vellend's (2010) suggestion to link to well-developed evolutionary theory.
Parameters	Variables or constants used to define and/or control the behavior or characteristics of a system or process.
Module	Small set of (possibly interacting) species/taxa/entities.
Statistical Model	A mathematical description of patterns observed in data. Many of them produce "metrics" that quantify

	patterns and allow for statistical evaluation.
Mechanistic Model	A simplified representation of an ecological system, constructed to understand, analyze, and predict outcomes within an ecological system such as a population, a community, or a metacommunity. Models are usually encoded in mathematical equations which track variables whose dynamics are modulated by parameters.

Appendix 2: How are communities and metacommunities represented by matrices?

When there are more than a few (say 3) species, using individual equations to solve their mathematical effects becomes difficult. Fortunately, mathematicians have developed tools that greatly facilitate this using matrix algebra. This applies to linked sets of equations, those that depend on the parallel structure of linear equations that are linked to each other because they have the same parameters. In community ecology this means taking a matrix that does this to the Lotka-Volterra models, then solving for a possible equilibrium point and linearizing the equations near that point to get something called “the Jacobian” of the matrix. It is then possible to solve for the dynamic behavior of these equations near the equilibrium. With this, the most important point is resolving whether this point is realistic and then stable/unstable. For an ecologist using Lotka-Volterra models, the first thing is to find out if the equilibrium point has any negative values; if so, this would mean negative densities and be meaningless. The second is that if it checks out (all positive values), then solving to find the conditions (if any) when the point is stable. This is done by calculating the ‘dominant eigenvalue’ that describes the tendency of the whole system to come back to equilibrium. Stability here is inferred if the ‘real part’ of this eigenvalue is less than 0.

May (1972) did this by developing the idea of the “community matrix”, which can be expressed as the Jacobian of the matrix describing the interactions among a group of species in a single local community (see Novak et al. for more details on this). This matrix is often denoted by the letter “A” and is the Jacobian of the interactions in the community.

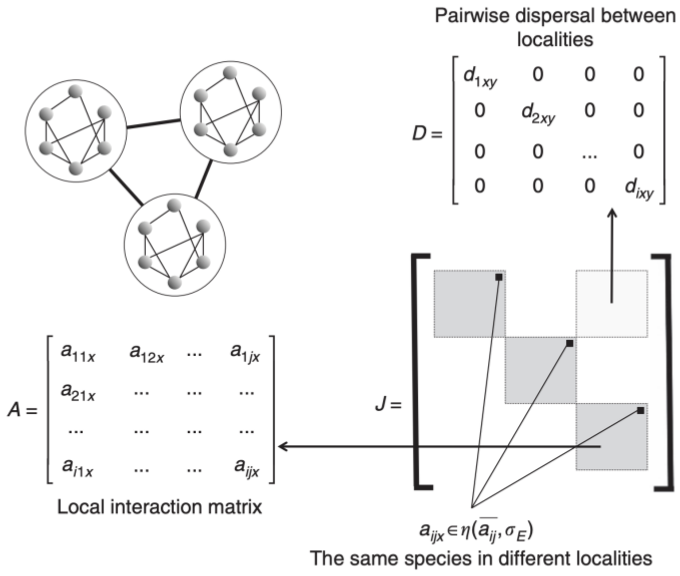

Figure B1: Going from local interaction matrices that describe how all species interact with each other (here this is the A matrix) by adding dispersal parameters to obtain the equivalent metacommunity matrix. This is done by lining up all the local matrices (keeping in mind that their separate values can be arbitrarily different) along the diagonal of the metacommunity matrix (darkly shaded part of the J matrix). Dispersal can then be described by other entries that are off diagonal light shading (matrices that have the structure shown in the D matrix. These matrices describe the dispersal rate (going from the column to the equivalent row elements of this matrix. They will only have entries along the main diagonal since species cannot change their identities while they disperse.

When May (1972) used this method on the A matrix where all off-diagonal entries were sampled randomly, he found that the system tended to have a very distinct stability only when:

$$\sigma \sqrt{c\,(S-1)} < m$$

Where S is the number of species present, c is the proportion of interactions that are not zero, δ is the variance in interspecific interaction strength, and m is the average intraspecific interaction strength.

Gravel et al. (2016) expanded this very basic model by asking how this would manifest in a metacommunity. To do so, they linked multiple local communities via dispersal as described in Figure B1. The stability criterion was then:

$$\sigma \sqrt{c\,(S-1)/N} < m$$

Where N was the ‘effective number of ecologically distinct’ local matrices. That is to say, May’s inequality was modified because ecologically distinct communities acted multiplicatively to enhance diversity in the metacommunity. This was also true in each local community since dispersal at these levels allowed them to do so in each locality as sink populations as well.

Of course, there are many other ways that multispecies ecology has used matrices such as the Jacobian of the community (and now the metacommunity), but the comparison of May (1972) and Gravel et al. (2016) illustrates some of the major features that allow the modeling of purely local interactions (May 1972), here modeled as a completely disordered system, and metacommunity (Gravel et al 2016) dynamics, here modeled as a partially structured extension of May’s (1972) work.

Appendix 3: Contrasting challenges for modeling approaches at the interface between patterns and processes

Figures in the main text present a simplified contrast between approaches to modeling communities, focusing mainly on the “bottom-up” and “top-down” perspectives. In fact, there are other important axes of variation between models, implying different methods and difficulties when trying to bridge the process-pattern gap.

To understand why different modeling approaches may encounter distinct challenges, we categorize them as follows:

1. **How the model is selected/tested against patterns**
    a. **Qualitative prediction:** The model predicts the presence or absence of some qualitative feature in the data, e.g. environment filtering suggests a positive correlation between phylogenetic relatedness and co-occurrence. This is often the case of simple theoretical scenarios such as the main metacommunity paradigms (Leibold et al. 2004), and one common challenge is finding ways to compare how strongly the evidence supports one paradigm over another.
    b. **Fitting:** Alternatively, a model can also generate a range of hypothetical outcomes, among which we identify the best match(es) to our observations by quantitative model fitting. One important challenge is the risk of mis-specifying the model, so that the best fit within its range is not the most appropriate picture of the world.
    c. **Cross-pattern consistency:** If the model is successful at predicting one pattern, does it correctly predict other patterns? As the number of congruent patterns increases, they each act to more strongly validate the model.
2. **How it represents processes**

   a. **As qualitative assumptions:** The model, explicitly or sometimes implicitly, makes a statement about the presence or importance of one or more processes (e.g. absence of dispersal, dominance of abiotic filtering). This is especially relevant when we want to consider scenarios that differ in many ways (e.g. aquatic vs terrestrial biomes, wind dispersal vs chemotaxis) so that they are more easily represented as distinct archetypes rather than on a very multidimensional continuum. The risk is not knowing how different assumptions would have changed the model's predictions.
   b. **As parameters:** Quantitative knobs relating to processes (e.g. rate of each process).This is especially tractable when we want to consider scenarios that differ continuously in a few significant ways. One challenge is interpreting parameter values and how they relate to predictions, for instance interaction rates and dispersal rates are in different units and cannot easily be compared to determine which of the two processes is most important, and theoretical analysis is needed to know how strong dispersal must be to homogenize a metacommunity.

3. **Level of description:**
   a. **Disaggregated**: We retain species identities by characterizing them by their species-specific parameters, and we seek to derive their individual contributions to overall community-level patterns. The usual problem in metacommunities is the lack of sufficient statistical power to resolve these properties for up to hundreds of species.
   b. **Aggregated:** We lose species-specific labels, and instead look at overall distributions or summary parameters and metrics. This method characterizes "disordered systems models" described in text. It can, however, produce meaningless results when we aggregate variables that cannot be (for instance, it is not obvious that we can sum biomasses or richness of species at different trophic levels).

      We note that both may be used together. For example, Guisan and Rahbek (2011) combine macroecological constraints and SDMs to obtain more informative predictions. Or we can fit a JSDM at the species level (as in Leibold et al. 2022) and then aggregate some components over all the species. We also note that, in principle, there can be some sort of intermediate level aggregation such as described for "partially structured models".

4. **Process interplay:** How does our model assume that different processes (e.g. environments, dispersal, interactions, including all the processes that make up the "context" e.g. biome type) come together?
    a. **Totally separable:** we can ask of each process type what is its contribution to patterns, irrespective of what others are doing (possibly after some transformation, e.g. GLMs and GAMs). For instance, an SDM attempts to quantitatively partition variance in species abundance into effects of spatial fluxes and environment.
    b. **Totally entangled:** Various specific combinations/rankings of processes can give various outcomes, without any way to separate them into contributions of each process. For instance, the "patch dynamics" metacommunity paradigm captures the effects of having dispersal and stochasticity and local interactions acting together, without being able to meaningfully separate these processes.
    c. **Dynamical models:** A special case where we assume that processes are totally separable in the short-term rate of change (often not measured), but entangled in outcomes such as observed abundances.

    d. **Multiple regimes** can be seen within the same model (often seen for dynamical models), several "regions" within which processes have consistent (and at least conceptually, if not practically/quantitatively, separable) effects on patterns/outcomes, but these effects differ from region to region – e.g. the notion of "limiting factor", with a region of limitation where increasing the factor steadily increases its importance and consequences, then a region without limitation where changing it has no impact.

5. **How a model is interpreted:** Do we try to use the model to infer underlying mechanisms and resolve causal relations that link process and pattern? (Our ability to do so is conditioned by the choices made along the previous axes.)
    a. **Prediction only:** Model assumptions and parameter values are not interpreted in any way, but only used to predict patterns in other similar situations. Processes thus only intervene in building the model (e.g. deciding which parameters and relations to include), but not in interpreting it. This includes the usual practice involving SDMs and JSDM to forecast future changes in species distributions under changing environmental or landscape conditions.
    b. **Qualitative agreement:** If the model correctly predicts patterns, we interpret this as suggesting that its qualitative assumptions regarding

processes are correct (for instance, seeing a spatial checkerboard pattern is interpreted as indicative of competition).

c. **Partitioning:** The values of model parameters are not necessarily meaningful in themselves, but they are used to deduce the relative importance of various processes e.g. by variance partitioning.

d. **Theory:** Model assumptions and parameter values must be analyzed theoretically to clarify their ecological implications (for instance, we have a theory of how strong interactions should be to induce nonequilibrium dynamics within a patch).

Some Examples:

We find that models linking processes and patterns span much of (if not all of) the range of attributes described above. To illustrate this, Table S1 classifies model attributes used across a wide array of studies, with many of these coming from our own work as individual researchers.

Model	**Selection by data/ patterns**	**Representation of processes**	**Level of description**	**Process interplay**	**Interpretation**
Competitive checkerboard vs null model	Presence or absence of specific pattern	Categorical: competition vs stochasticity	Species	Single dominant process	Qualitative
Functional trait over/underdispersion	Presence or absence of specific pattern	Categorical: competition vs env. selection	Species	Single dominant process	Qualitative
Biodiversity-Function relations	Fitting a specific pattern	Parameters: selection/complementarity	Community	Separable	Partitioning
Environmental niche model	Fitting to full compositional data	Parameters	Species	Separable	Partitioning
Joint Species Distribution	Fitting to full compositio	Parameters	Species	Separable	Partitioning

Model	nal data				
Fitted gLV	Fitting to full composition or time series data	Parameters	Species	Dynamical, multi-regime	Requires theory
Probabilistic macroecology	Fitting several distributions	Categorical: species fitness + stochasticity	Community	Separable	Qualitative
Disordered systems model	Global means and variances (covariances) of parameters	Parameters	Community	Separable	Requires theory
Partially structured model	Subsetted means and variances (covariances) of parameters	Parameters	Community	Separable	Requires theory