Scale Both Confounds and Informs Characterization of Species Coexistence in Empirical Systems

Adam Clark,[1,2,3,*] Helmut Hillebrand,[2,4] and W. Stanley Harpole[1,2,5]

1. Department of Physiological Diversity, Helmholtz Centre for Environmental Research (UFZ), Permoserstrasse 15, 04318 Leipzig, Germany; 2. German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Deutscher Platz 5e, 04103 Leipzig, Germany; 3. Synthesis Centre for Biodiversity Sciences (sDiv), Deutscher Platz 5e, 04103 Leipzig, Germany; 4. Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University Oldenburg, Schleusenstraße 1, 26382 Wilhelmshaven, Germany; and Helmholtz Institute for Functional Marine Biodiversity at the University Oldenburg (HIFMB), Ammerländer Heerstrasse 231, 26129 Oldenburg, Germany; 5. Institute of Biology, Martin Luther University Halle-Wittenberg, Am Kirchtor 1, 06108 Halle (Saale), Germany

ABSTRACT: Identifying stable coexistence in empirical systems is notoriously difficult. Here, we show how spatiotemporal structure and complex system dynamics can confound two commonly used stability metrics in empirical contexts: response to perturbation and invasion rate when rare. We use these metrics to characterize stable coexistence across a range of spatial and temporal scales for five simulated models in which the ability of species to coexist in the long term is known a priori and for an empirical old field successional time series. We term the resulting multivariate distribution of metrics a "stability fingerprint." In accordance with a wide range of classic and recent studies, our results demonstrate that no combination of empirically tractable metrics or measurements is guaranteed to "correctly" characterize coexistence. However, we also find that heuristic information from the stability fingerprint can be used to broadly characterize dynamic behavior and identify circumstances under which particular combinations of species are likely to persist. Moreover, stability fingerprints appear to be particularly well suited for matching potential theoretical models to observed dynamics. These findings suggest that it may be prudent to shift the focus of empirical stability analysis away from quantifying single measures of stability and toward more heuristic, multivariate characterizations of community dynamics.

Keywords: population stability, community stability, spatial scale, temporal scale, perturbation, mutual invasibility.

Introduction

The primary purpose of this article is to demonstrate how an empiricist might study stable coexistence in real-world settings. Such assessments are of great practical importance, as they are necessary prerequisites for prescriptive manage-

* Corresponding author; email: adam.tclark@gmail.com.
   ORCIDs: Clark, https://orcid.org/0000-0002-8843-3278; Hillebrand, https://orcid.org/0000-0001-7449-1613; Harpole, https://orcid.org/0000-0002-3404-9174.

ment of ecological systems and for testing theory (Levin 1992; Murdoch 1994; Ives and Carpenter 2007). However, the proper application of existing coexistence metrics requires substantial a priori knowledge about community structure, or measurements that are infeasible in empirical contexts (Stommel 1963; Turelli 1986; Donohue et al. 2016). Consequently, predictions about whether a particular combination of species is likely to coexist in the long term remain rare, particularly in diverse, real-world systems (Lawton 1999; Chesson 2003; Siepielski and McPeek 2010; Adler et al. 2013).

Although there is no universally accepted definition of "stable coexistence," most ecologists would probably agree that a stably coexisting community should retain a particular combination of species despite minor disturbances, even when measured at some point in the distant future. This property is called "persistence," and it encompasses a wide range of dynamic behaviors (Pimm 1984; Anderson et al. 1992; Grimm and Wissel 1997). Persistent states might be centered around a fixed point (Saavedra et al. 2017), an oscillatory cycle (May 1974), a stochastic distribution (Turelli 1986), or a moving trend (Chesson 2017). They might be robust to infinitely small perturbations (May 1973), perturbations that fall within a fixed range (Armstrong and McGehee 1980), or any perturbation of any size (MacArthur 1970). Species might theoretically be able to persist indefinitely or only up to some finite time horizon (Turelli 1980).

In practice, ecologists typically classify coexistence using one of two methods. Perturbation tests measure species responses to small disturbances (i.e., asymptotic stability; May 1973; Arnoldi et al. 2016). Stable coexistence is indicated if all species return to their predisturbance state (which could be a fixed abundance or a dynamic trajectory). Alternatively, invasion analyses test whether species can reinvade a community after being driven to low abundance (i.e., mutual invasibility; Chesson 2000b, 2003). If all species can invade, this

is taken to imply stable coexistence. These two metrics are conceptually similar, as they measure, respectively, the behavior of a community as it is pushed away from equilibrium toward a boundary or away from a boundary toward equilibrium.

When applied correctly, either metric can rigorously codify opportunities for stable coexistence. However, fulfilling the necessary criteria can be onerous. For example, because perturbation tests only assess local stability (i.e., in the vicinity of individual equilibria), every possible equilibrium must be tested separately (Anderson et al. 1992; Saavedra et al. 2017). Similarly, in systems with multiple stable states—as might occur due to Allee effects, trophic structure, or intransitive competition—species may be able to coexist at high abundance even if they cannot invade when rare (Barabás et al. 2018). Thus, invasion analyses typically need to be augmented (Turelli 1981, 1986; Chesson and Ellner 1989), for example, by introducing species at high enough density to exceed Allee thresholds (Chesson 2000*b*, pp. 359–360), jointly introducing predator/prey or mutualist pairs (Chesson and Kuang 2008; Levine et al. 2017) or assessing invasions from every possible boundary point (Chesson 2018, p. 1786).

For several reasons, it seems unlikely that these criteria can be fully satisfied in empirical systems. First, without perfect a priori understanding of system structure, it is unclear which test augmentations might be necessary. For example, if a species fails to invade, this could simply indicate that Allee thresholds were not overcome or that an obligate mutualist was omitted (Levine et al. 2017; Barabás et al. 2018). Likewise, if a perturbed species does not recover its initial state, this might indicate lack of stability or that the equilibrium is nonstationary (Chesson 2017). A second challenge is that real-world systems are subject to demographic stochasticity, which can obscure effects of perturbations or invasions (Durrett and Levin 1994). For example, failure to recover after a disturbance or failure to invade from low density might be indicative of random chance rather than average system behavior (Turelli 1981, 1986; Tilman 2004). Finally, even in well-understood systems, it may not be feasible to conduct all necessary measurements. For example, diverse communities have enormous numbers of potential equilibria, each of which might need to be tested in separate experiments (Levine et al. 2017; Saavedra et al. 2017; Chesson 2018). This challenge is compounded by effects of spatial and temporal structure, which can cause community responses to perturbations and invasions to vary greatly across scales (Stommel 1963; Levin 1992). Worse still, the scales required for coexistence to manifest may differ across species (Leibold and Chase 2018) and "may be much larger than is considered in most models and field studies" (Chesson 2000*b*, p. 344), for example, continents or centuries (Lawton 1999; Davis and Shaw 2001; Ricklefs 2008).

These fundamental challenges are potentially discouraging, and, in the words of Robert May, lead to "rather gloomy thoughts as to the extent to which one can, or cannot, hope to give the empiricist some precise, measurable definition of stability" (May 1973, p. 213). In other words, it seems likely that in empirical systems, no feasible combination of additional data or improved methods can conclusively demonstrate stable coexistence. In response, we suggest a shift of focus—away from rigidly interpreted tests and toward a more heuristic approach. In doing so, we are inspired by classical insights from theory (Turelli 1986; Levin 1992; Lawton 1999) and from a growing consensus in resilience ecology, which suggests that stability is best summarized using suites of metrics (Pimm 1984; Grimm and Wissel 1997; Carpenter et al. 2001; Ives and Carpenter 2007; Donohue et al. 2013, 2016; Arnoldi et al. 2016; Hillebrand et al. 2018; Meyer et al. 2018; Zelnik et al. 2018). Building on these results, we demonstrate (i) that empirically feasible measurements of coexistence reveal complex stability landscapes, which confound attempts to classify stability in a definitive manner, but (ii) that by combining information from across multiple metrics and scales of measurement, this complexity can be used to inform studies of general system behavior.

Methods

We approach our analyses from the perspective of an empiricist who has been confronted with a novel ecosystem. Given limited ability to conduct measurements and no a priori information about its constituent species, we attempt to characterize coexistence. Due to these constraints, our methods are never guaranteed to "correctly" identify stable coexistence. We therefore follow the advice of Turelli (1986, p. 331), who, in reference to applying invasion analyses to complex community dynamics, suggests, "In desperation, one can ignore mathematical rigor, apply a heuristic coexistence criterion, then do simulations to check its accuracy and hope for the best." In other words, in cases where no theoretically sound, empirically tractable metrics exist, one must make do with imperfect, practical alternatives. In this spirit, we optimistically present a heuristic solution that provides useful insights to the problem of characterizing species coexistence in ecological communities.

We proceed in three stages. First, we introduce five spatially and temporally explicit models of ecological communities. It is known a priori from theory that in some of these models long-term coexistence is possible, whereas in others it is not. However, we hide this information from the empiricist. Second, we allow the empiricist to conduct perturbation tests and invasion analyses in these systems across a variety of spatial and temporal extents. These do not necessarily correspond to the "best" scales for measurements but rather represent subsets of scales to which measurements might be limited due to practical constraints (fig. 1). We term the resulting multiscale distributions of metrics "stability fingerprints"
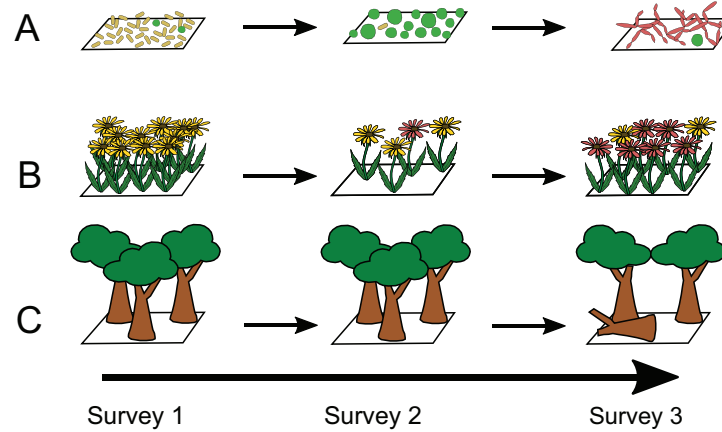
Figure 1: Because of a combination of physical limitations, financial barriers, and historical legacy, the spatial and temporal scales at which one can reasonably observe and manipulate ecological systems do not always correspond to those that are useful for studying coexistence. Consider a time series comprised of three observations, in each of three different types of communities. *A*, Samples of microbial communities might include millions of individuals, and hundreds of generations can pass between surveys. *B*, Herbaceous plant communities can contain hundreds of individuals per square meter, and measurements of species-level biomass are sufficiently destructive and time-consuming that only a small fraction of the total landscape can be surveyed. *C*, For tree communities, survey plots might contain only a few individuals, but dynamics can be very slow, playing out over decades or centuries.

because they summarize unique information about system dynamics, and we use them to assess stability in each model. Last, we apply this fingerprinting procedure to characterize empirical dynamics during old field succession, using data from a 90-year chronosequence. This worked example serves as a guide for implementing our approach in a system where very few of the theoretical requirements for stability analysis have been met.

Simulating Model Dynamics

Our models are individual-based stochastic patch-occupancy simulations with similar structures. For all but one model (plant-soil feedback; see below), we implement simulations using Gillespie's method following Lehman et al. (2012). This method has several advantages, including that demographic stochasticity arises as an emergent property and that simulations perfectly match analytically tractable models at large spatial scales (i.e., master equations, sensu Black and McKane 2012), which contributes a priori theoretical knowledge about their stability. Simulations occur on a 100-by-100-site grid, where sites can be occupied by at most a single individual. For each species i, dynamics depend on interspecific interactions and stochastic colonization and mortality events, with average rates c_i and m_i, respectively. For simplicity, we assume uniform dispersal across all patches. See sections A1 and A2 of appendix A for model derivations and theoretical stability properties, sections B1 and B2 of appendix B for details on Gillespie's method and for results in the absence of

demographic stochasticity, and appendix E for source code (apps. A–E are available online).[1]

The first model, Levins, is a spatially and temporally explicit realization of the metapopulation model of Tilman (1994) and Levins (1969). At large spatial extents (i.e., large grids), average dynamics of species considered across all patches (i.e., the maximum spatial extent) follow Tilman (1994):

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = c_i p_i \left(1 - \sum_{j \leq i} p_j\right) - m_i p_i - p_i \sum_{j < i} c_j p_j, \quad (1)$$

where p_i is the fraction of sites occupied by species i. Competition in the Levins model is perfectly transitive, such that species $i = 1$ is competitively superior to all other species, species $i = 2$ is competitively superior to all species but species 1, and so on. The first term in equation (1) represents colonization into sites that are empty or occupied by inferior competitors, the second term represents mortality, and the third term represents displacement caused by superior competitors. Following Tilman (1994), species equilibrium abundance p_i^* can be calculated sequentially starting with the best competitor as

$$p_i^* = 1 - \frac{m_i}{c_i} - \sum_{j < i} p_j^* \left(1 + \frac{c_j}{c_i}\right) \quad (2)$$

if $p_i^* > 0$ and 0 otherwise.

---

1. Code that appears in *The American Naturalist* is provided as a convenience to readers. It has not necessarily been tested as part of peer review.

For species i to persist at equilibrium (i.e., $p_i^* > 0$), c_i must exceed m_i, and colonization rates of inferior competitors must be sufficiently large to offset competition. Coexistence is globally stable, meaning that species are drawn toward the equilibria in equation (2) from any nonzero starting abundance (fig. 2a, 2b).

Our second model, disturbance, is almost identical to the Levins model but also includes periodic events every D time steps that destroy fraction d_i individuals of each species. When considered across time (i.e., periods with and without disturbances), disturbances increase average mortality rate by a factor $\log(1 - d_i)/D$. Time-averaged abundances and global stability at large spatial scales can be calculated using these average rates and equation (2), just as in the Levins model (Chesson and Huntly 1997; Barabás and Ostling 2013; see sec. A2 of app. A and fig. S1 for more details;
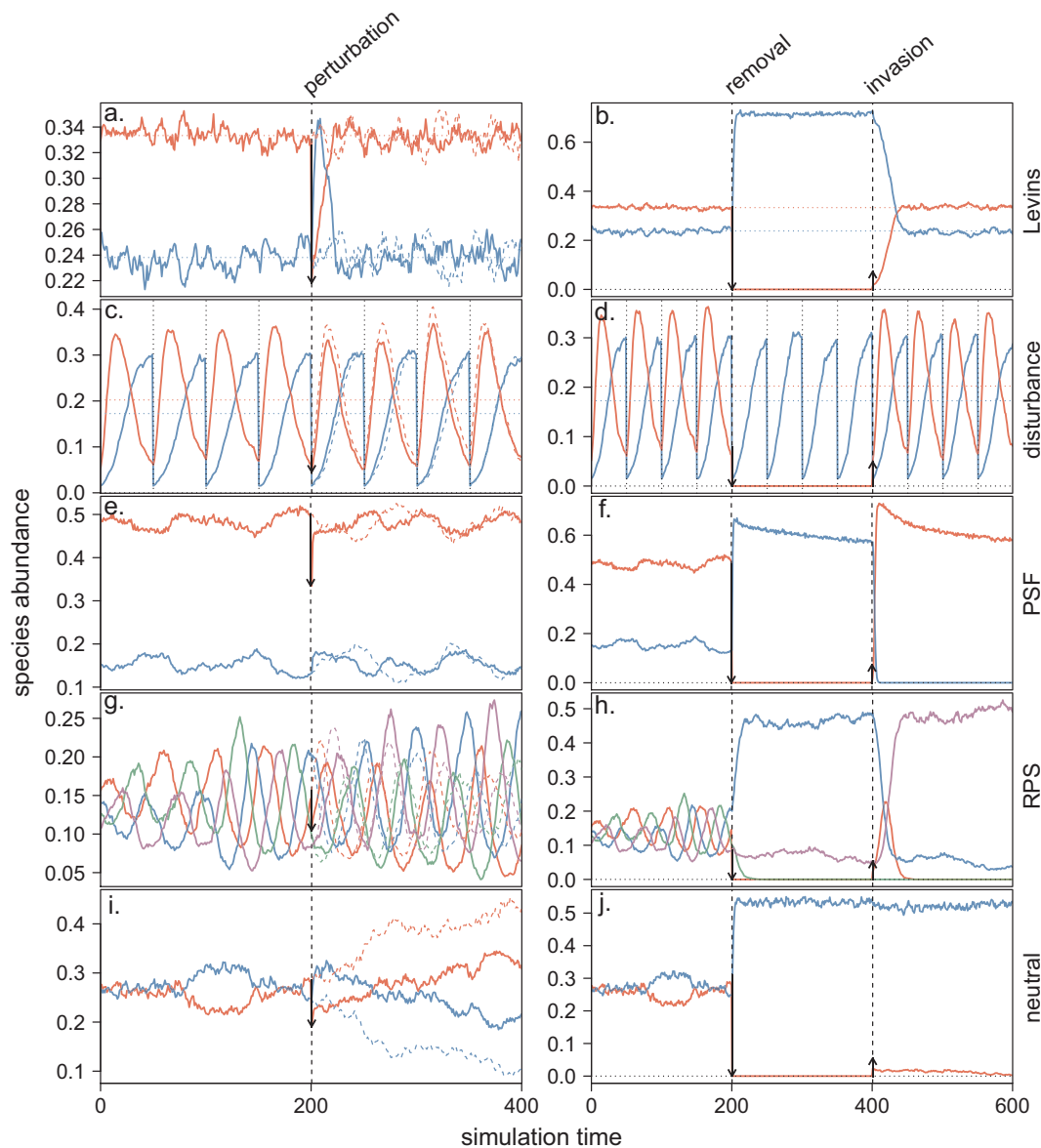
Figure 2: Example dynamics of each model at the maximum spatial extent (i.e., across all simulated sites). Each row shows results for a single model. Colored lines show species abundances. Vertical dashed lines and arrows show perturbation events for perturbation tests (*a, c, e, g, i*) or removals and invasion events for invasion test (*b, d, f, h, j*), always for the red species. Colored dashed lines in *a, c, e, g,* and *i* show trajectory in the absence of perturbation. Horizontal lines in *a* and *b* show equilibria from equation (2), and those in *c* and *d* show mean abundances predicted from temporally averaged mortality rates (see sec. A2 of app. A for details). Vertical dotted lines in *c* and *d* show disturbance events.

figs. S1–S16 are available online). However, the more erratic dynamics caused by disturbances mean that short-term responses to perturbations—and therefore empirical estimates of stability metrics—can differ greatly between the Levins and the disturbance model.

Our third model, plant-soil feedback (PSF), follows a previously published model (Suding et al. 2013) in which spatial structure and demography are similar to those in the Levins model, but species alter the environment in patches they occupy. These effects build up over time, such that species either increase their own mortality and decrease that of their competitors or vice versa (for negative or positive feedbacks, respectively). Because of these feedbacks, we could not implement the PSF model using Gillespie's method (see sec. B1 of app. B for details). Given uniform dispersal, long-term coexistence requires negative feedbacks for all species (Suding et al. 2013). This coexistence is locally stable, such that communities can recover from small perturbations but large perturbations lead to alternate stable states. These alternate states form because negative feedbacks reduce species performance over time, eventually allowing invaders to displace resident species (fig. 2*e*, 2*f*; see sec. A2 of app. A and figs. S2 and S3 for details).

Our fourth model, rock-paper-scissors (RPS), has the same structure as the Levins model except that species follow an intransitive competitive hierarchy, such that in terms of competitive ability, species $i > j$, $j > k$, $k > l$, and $l > i$. All other species pairs can displace one another, but without hierarchical advantages. We consider a community of four species in which species share the same demographic rates m and c, which leads to neutrally stable oscillations around a fixed point (sensu Allesina and Levine 2011; Grilli et al. 2017). Demographic stochasticity therefore causes species to drift toward extinction, meaning that long-term coexistence of all four species is not possible (fig. 2*g*, 2*h*). However, because oscillations are compensatory across species, total summed community abundance acts like a single species in the Levins model and is globally stable around

$$\sum_{i=1}^{n} p_i^* = 1 - \frac{m}{c}. \tag{3}$$

Last, our fifth model, neutral, again follows the same form as the Levins model except that all species share the same demographic rates m and c and no species are competitively dominant (i.e., no species can colonize an occupied site). Population-level dynamics are therefore dominated by ecological drift sensu Hubbell (2001), and species cannot coexist in the long term (fig. 2*i*, 2*j*). As with the RPS model, equilibrium total community abundance is globally stable and is approximated following equation (3).

For simplicity, we simulated communities of two species for all models except for RPS, which included four species.

Whenever possible, we chose parameters that resulted in long-term persistence of all species (see sec. A3 of app. A for parameter values). For the Levins model, we included a fast-dispersing inferior competitor and a slow-dispersing superior competitor, for which globally stable coexistence is predicted by equation (2). For the disturbance model, we chose parameter values such that competitive exclusion was predicted in the absence of disturbance, but long-term coexistence was possible due to trade-offs between competitive ability and resistance to disturbance. For the PSF model, we included two species, both with negative feedbacks, which allowed locally stable coexistence. Recall that long-term coexistence is not possible in the RPS or neutral model.

Testing for Stable Coexistence

For each model, we characterized coexistence across a range of spatial and temporal extents (denoted s and t, respectively) by conducting perturbation tests and invasion analyses. These extents represent contiguous units of space and time, as might be sampled in an empirical study (e.g., a 1-by-1-m plot, measured over five field seasons). We considered extents ranging from 0.5% to 100% of the maximum spatial extent (i.e., all 100-by-100 sites) and from 1 to 200 simulated time steps.

We conducted manipulations and measurements only at the focal extents corresponding to each test. These procedures were meant to mimic plot-based ecological observations, where only a fraction of the landscape can be manipulated and observed (e.g., fig. 1; see sec. A4 of app. A for details). We simulated 20,000 iterations of each model and report the median test result across iterations (see fig. S4 for distributions). This procedure generated fingerprints for each model that summarized stability across extents.

To conduct perturbation tests, we applied small perturbations to a species (20% reduction in abundance) and compared dynamics to those in simulations without the perturbation (e.g., as might be accomplished by comparing control and treatment plots; see fig. 2, perturbation). We quantified this response as the average rate of return:

$$r_e(s, t) =$$
$$\text{max(over all species } i) \left[\frac{1}{t} \log \left(\frac{|p_{i,s,t} - p_{i,s,t}^0|}{|p_{i,s,0} - p_{i,s,0}^0|} \right) \right], \tag{4}$$

where $p_{i,s,t}$ describes the abundance of species i measured at spatial extent s and temporal extent t and $p_{i,s,t}^0$ describes what the abundance of species i would have been had the perturbation not occurred. If the system returns toward its initial state, then $r_e < 0$, which is interpreted as stability. Thus, r_e represents the most positive—and therefore the least stable—response observed across all species (see fig. S5*a*–S5*c* for examples) and is formally an approximation of the Lyapunov exponent (see sec. A5 of app. A for details and justification).

To conduct invasion analyses, we removed a species from the community, allowed the resulting community to equilibrate to account for transient dynamics, and then reintroduced the species at low abundance (5% of unoccupied sites) sensu Chesson (2000*b*) (e.g., as might be accomplished through experimental introduction of species into an existing community; see fig. 2, removal and invasion). We quantified invasions using the average growth rate:

$$r_0(s, t) =$$
$$\min(\text{over all species } i) \left[\frac{1}{t} \log \left(\frac{p_{i,s,t}}{p_{i,s,0}} \right) \right], \qquad (5)$$

where t describes time since reinvasion, and r_0 describes the minimum rate observed across all species (i.e., the least positive, and therefore the least stable; see fig. S5*d*, S5*e* for examples).

Note that if systems recover from perturbations or invasions, both r_e and r_0 necessarily approach zero for large t. Ideally, this property should be controlled for by considering different subsets of temporal extents (Sheil and May 1996). However, this may not always be possible in empirical settings (e.g., fig. 1*a*, where system dynamics are fast relative to measurements). Thus, particularly for large t, our results represent potential outcomes that might arise due to sampling constraints rather than ideal tests that are best suited for detecting coexistence.

Empirical Example

To demonstrate how our methods might be applied to real-world systems, we analyzed a 90-year-old field successional chronosequence from the Cedar Creek Ecosystem Science Reserve in Minnesota (Inouye et al. 1987; Clark et al. 2019). This chronosequence includes 23 fields, abandoned from agricultural use between 1927 and 2015. In each field, species-level percent cover of herbaceous plants has been surveyed in 100 permanent 0.5-by-1-m plots, roughly every 5 years since 1983. We chose this data set because of its uniquely large range of measured spatial and temporal extents (see fig. S6). Note, however, that very few theoretical assumptions for coexistence analysis are met in this study; for example, all species are perturbed to low abundance simultaneously through tilling, and no control observations in undisturbed fields are available. Data can be accessed on the Long Term Ecological Research (LTER) network data portal (https://doi.org/10.6073/pasta/aa029df8f7a6091ea879ceb5c6673963; Knops 2018).

We combined species into functional groups following Clark (2017) and retained the three most abundant categories: annual species, cool-season (C_3) perennial grasses, and warm-season (C_4) perennial grasses/sedges. These accounted for >80% of total cover. Importantly, successional dynamics and long-term persistence for these groups are well known:

annuals are primarily found early in succession, C_3 grasses are typically midsuccessional, and C_4 grasses dominate late in succession (Clark et al. 2019). See appendix C for more details on the site and analysis methods.

We conducted analyses in three steps. First, to measure stability we treated succession as an invasion event and approximated r_0 based on changes in percent cover over successional time following equation (5). Similar metrics might be calculated following any large disturbance that reduces species to low abundance—for example, experimental manipulations or natural events. Although we could not do so here, in other systems it may be possible to approximate r_e based on differences between perturbed and unperturbed replicates, following equation (4).

Next, we calculated r_0 across observed spatial and temporal extents to generate an empirical stability fingerprint. We used years since agricultural abandonment as a proxy for t, yielding extents ranging from 1 to 89 years after disturbance. For spatial extents, we aggregated nested subsets of plots based on proximity (e.g., neighboring plots) and calculated s as total surveyed area, ranging from 0.5 to 406 m^2 (n.b.: we excluded some experimentally burned plots from analyses). Similar methods could be employed in any system with repeated sampling through time or spatially replicated observations. Unlike our analyses of simulated data, we present rates scaled by time (i.e., $r_0 t$), as this helps visualize dynamics over long temporal extents (n.b.: this transformation does not influence subsequent analyses of metric distributions or model comparisons, which focus on relative values within individual spatiotemporal extents). In general, we suggest similar transformations for any analysis of long-term data, especially if systems appear to approach an equilibrium.

Finally, we compared empirical fingerprints from the old fields to those from two models: (i) Levins-OF, a realization of the Levins model parameterized with demographic rates following Clark (2017); and (ii) neutral-OF, a realization of the neutral model based on species average demographic rates. We test the Levins-OF model because it has been hypothesized as a model of successional dynamics at Cedar Creek (Tilman 1994; Clark et al. 2019). We include the neutral-OF model as a simple alternate model because it requires little a priori information for parameterization. To emulate the old fields' disturbance history, we initialized models using abundances observed immediately after abandonment and simulated 90 years of succession. We then compared the empirical and simulated systems by calculating the likelihood of empirical fingerprints given simulated fingerprints (see app. D for details).

Results

Model Dynamics

The Levins model most clearly demonstrated properties that are commonly associated with stable coexistence. At large

extents, species recovered rapidly following perturbations and invasions and equilibrated around analytical expectations from equation (2) (fig. 2*a*, 2*b*). Results were similar for the disturbance model, except that species followed oscillatory cycles (fig. 2*c*, 2*d*). For the PSF model, species recovered from perturbations, but invasions led to alternate stable states wherein each species could displace its competitor but the two could not coexist (fig. 2*e*, 2*f*). For the RPS model, perturbation tests led to persistent changes in oscillatory dynamics (fig. 2*g*; n.b.: solid and dashed lines do not converge). In invasion tests, removing one species led to extinction of a second species, which left the removed species' superior competitor unchecked (fig. 2*h*). Thus, subsequent reinvasion was only transiently successful. Last, as expected for the neutral model, both perturbation and invasion tests changed system dynamics, with no recovery to initial state (fig. 2*i*, 2*j*).

Stability Fingerprints

As expected, neither heuristic metric "correctly" identified stable coexistence across all models and scales (recall that long-term coexistence is possible in the Levins, disturbance, and PSF models but not in the RPS and neutral models). Across models, r_e generally increased with spatial extent, such that below 1% of the maximum extent it was always positive (fig. 3*a*–3*e*). For r_0 patterns were more variable, although it often increased at larger spatial extents (fig. 3*f*–3*j*). Both metrics approached zero for large temporal extents, as expected for time-averaged growth rates (Sheil and May 1996).

For the Levins model, both metrics correctly indicated that long-term coexistence was possible with negative r_e (i.e., recovery from perturbation) and positive r_0 (i.e., successful invasion), except at extreme spatial or temporal extents (fig. 3*a*, 3*f*). For the disturbance model, metrics gave conflicting results. Although long-term coexistence was possible, immediately after disturbances we found negative r_e (indicating stability; note that narrow bands are visible at larger spatial scales) and negative r_0 (indicating lack of stability), whereas between disturbance events we detected positive r_e and positive r_0 (fig. 3*b*, 3*g*). For the PSF model, for which long-term coexistence is locally stable, results were similar to those for the Levins model, despite the existence of alternative stable states (fig. 3*c*, 3*h*). Finally, for the RPS and neutral models, both metrics correctly indicated lack of stability at large spatiotemporal extents (fig. 3*d*, 3*e*, 3*i*–3*j*). However, for both models r_0 incorrectly indicated stability at small spatial extents, and for the RPS model r_0 also indicated stability during periods of transient reinvasion.

Empirical Example

Although our empirical analysis of successional dynamics failed to meet many theoretical requirements, results nevertheless accorded with species hypothesized successional niches. Across spatial extents, annuals showed successful invasions for roughly the first 10 years of succession, followed by population declines over longer temporal extents (fig. 4*a*). For C_3 grasses, growth was generally negative for the first 5 years and then positive for longer temporal extents, although growth was always positive at spatial extents above ~25 m² and declined somewhat at temporal extents above 20 years (fig. 4*b*). Trends for C_4 grasses were similar, except that positive growth only began around successional year 20 or spatial extents above 70 m² and then increased monotonically for larger temporal extents (fig. 4*c*).

Correspondence between simulated and empirical fingerprints was almost always higher for the Levins-OF model than for the neutral-OF model, particularly for temporal extents above 10 years and spatial extents above 50 patches (fig. 4*d*–4*f*). Unlike the neutral-OF model, the Levins-OF model successfully predicted declines in r_0 with field age for annuals and increases in r_0 with field age at smaller spatial extents for C_3 and C_4 grasses (fig. 4*g*–4*i*). Interestingly, this was true even though both models predicted similar dynamics over the first decade of succession and consistently overpredicted abundances (see fig. S7).

Discussion

Across all models, no combination of empirically tractable measurements was guaranteed to correctly identify stable coexistence. While not surprising, this result serves as a reminder that attempts to characterize coexistence are necessarily heuristic from the perspective of an empiricist—that is, given limited ability to make measurements and imperfect understanding of system structure (Turelli 1986; Levin 1992; Murdoch 1994; Lawton 1999; Donohue et al. 2016). Thus, definitive statements about coexistence in empirical contexts are probably neither testable nor warranted.

More encouragingly, our results show that multimetric, multiscale empirical tests can still produce useful results. Specifically, multivariate stability fingerprints contribute three main types of insight, described in detail below. First, based on empirically tractable measurements, fingerprints summarize major components of system dynamics, which often relate closely to long-term coexistence. Second, in cases where predictions about long-term dynamics are unclear or confounded, fingerprints help identify potential drivers of uncertainty. Third, fingerprints aid in identifying potential mechanisms underlying observed dynamics. Jointly, these findings support conjectures from a broad range of studies, which suggest that the focus of stability analysis should be shifted toward more holistic, multivariate assessments (Levin 1992; Grimm and Wissel 1997; Ives and Carpenter 2007; Donohue et al. 2016; Zelnik et al. 2018).
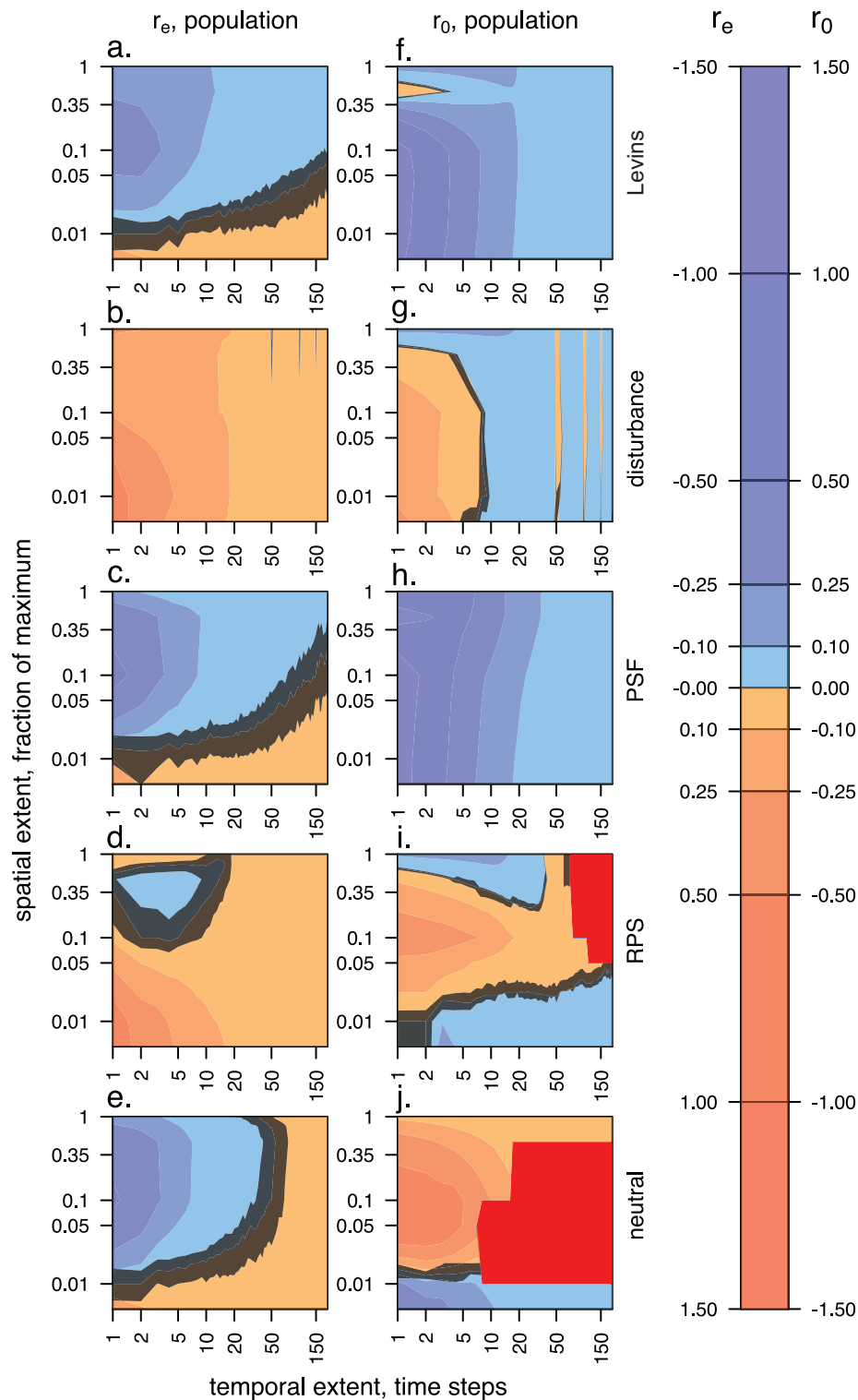
Figure 3: Stability fingerprints summarizing median results across spatial and temporal extents, based on 20,000 iterations of each model. Columns show different metrics, and rows show different models. Colors show the median value for maximum r_e (*a–e*) or minimum r_0 (*f–j*) observed across species. Cooler colors indicate greater measured stability (recall that negative r_e or positive r_0 indicate stable coexistence). Bright red regions indicate extinction of the manipulated species in more than one-third of simulations. Dark gray shading shows regions where between-simulation variability is sufficiently large that >50 observations would be required to detect a significant difference from zero. Shading is semitransparent, so that the sign of the index is still visible. For more information on the distribution of results, see figure S4.
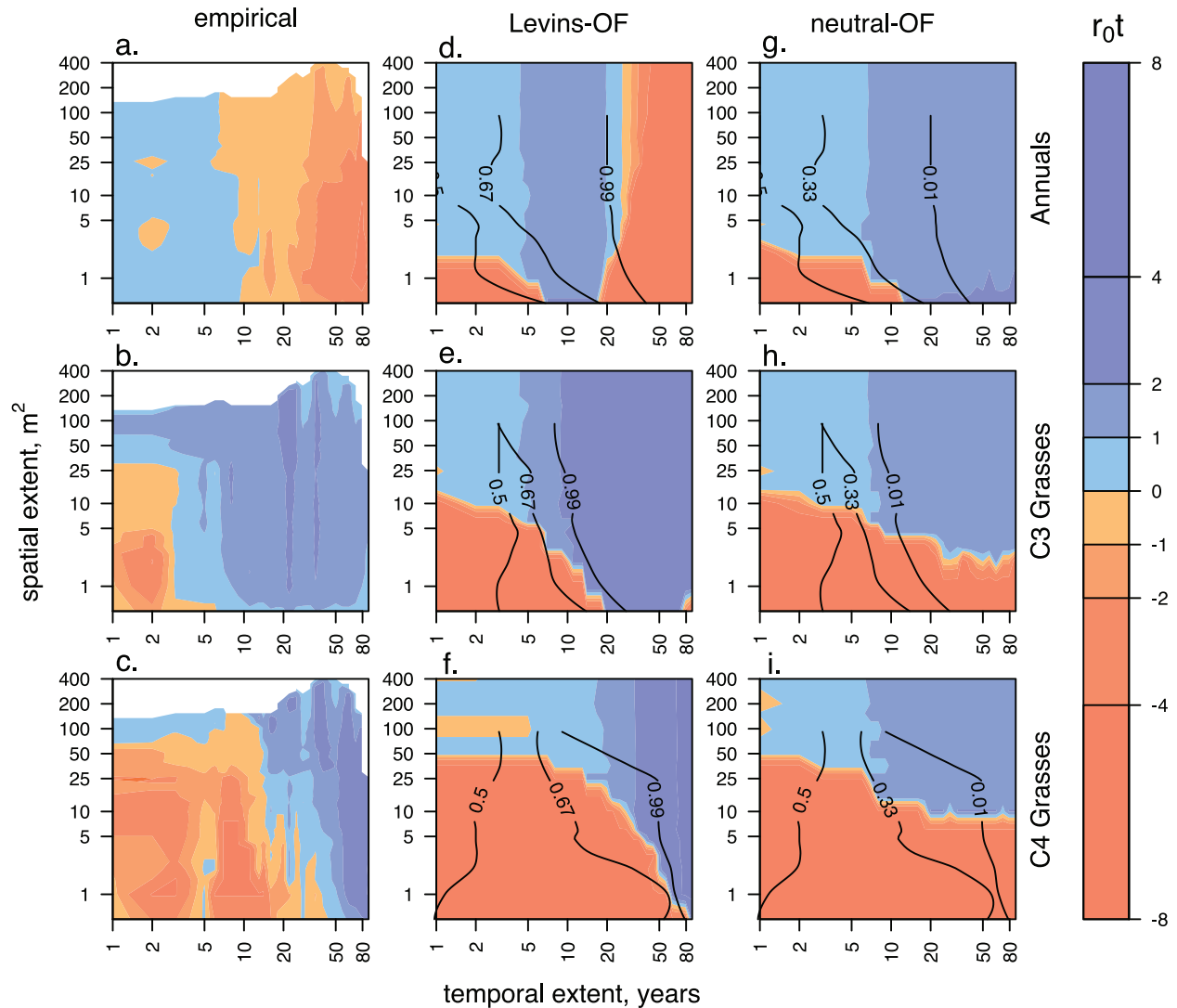
Figure 4: Stability fingerprints for successional dynamics in old fields at Cedar Creek generated from approximated invasion statistic r_0. Rows show results for different species groups, and columns show fingerprints for observed dynamics (*a–c*), simulated results from the Levins-OF model (*d–f*), and simulated results from the neutral-OF model (*g–i*). Colors show median stability metrics, either from 20,000 bootstrapped samples drawn from old field surveys (*a–c*) or 20,000 simulations (*d–i*). Contour lines show relative cumulative likelihood of observed pattern given the Levins-OF (*d–f*) versus neutral-OF (*g–i*) models (defined as $\mathrm{RCL}_{m|\mathrm{OF}}$; see app. D and fig. S16 for more details). Note that unlike equation (5) and figure 3*f*–3*j*, r_0 is multiplied by temporal extent t to make patterns at larger temporal extents clearer (Sheil and May 1996).

Applying Fingerprinting

Although the simulation methods we employ are complex, note that generating the stability fingerprints themselves is relatively simple: r_0 is effectively a log response ratio comparing population size at time t versus initial observed population size, and r_e compares the log ratio between two observed time series (e.g., replicates with and without an experimental perturbation). Moreover, our results do not rely on a theoretically "optimal" perturbation. For example, our analysis of

empirical old field data is in many ways a worst-case scenario for stability analyses, as all species were perturbed to low abundance simultaneously and no control plots were available.

Despite their simplicity, fingerprints generally succeeded in characterizing complex dynamics. For example, matching a priori expectations for our models, r_0 demonstrates that reinvasion in the RPS model is transient (fig. 3*i*), and r_e shows that in the neutral model species abundances can recover from small perturbations, but only temporarily (fig. 3*e*; Hubbell 2001; Grilli et al. 2017). Likewise, for empirical old field

dynamics, r_0 shows that growth rates for C_3 and C_4 grasses are positive only when measured at extents above a few dozen plots or temporal extents of greater than 5–10 years (fig. 4a–4c). This result accords with theoretical expectations for the Levins model, which suggests that long-term persistence of late-successional species requires large spatial extents (Tilman 1994).

Note that stability fingerprints are strongly influenced by perturbation type. We demonstrate two types of responses, which effectively represent different ends of a continuum ranging from small and instantaneous (r_e) to large and distributed across time (r_0) (Ives and Carpenter 2007). However, many other perturbation types also contain useful information. For example, although the regularly occurring mortality events in the disturbance model generally confounded our tests, species responses to these events help demonstrate the long-term viability of coexistence (see fig. S8). Combining information from across many species or many different perturbation types, including experimental manipulations and natural events, could therefore greatly increase information availability, especially in systems where observational scales are limited (Pimm 1984; Carpenter et al. 2001; Donohue et al. 2013, 2016; Arnoldi et al. 2018). This approach has recently been termed "probing" and suggested as a general method for classifying complex dynamics (Zelnik et al. 2018).

Addressing Problems

It is tempting to imagine that deviations between our predictions from individual indices and the actual long-term persistence of species are idiosyncratic to the metrics and models we consider. However, as noted, our results accord with a broad array of studies, which suggests that many of the mischaracterizations that we demonstrate are indicative of fundamental obstacles and unavoidable trade-offs (Turelli 1986; Levin 1992; Murdoch 1994; Chesson 2000a; Carpenter et al. 2001; Hubbell 2001; Levine et al. 2017; Saavedra et al. 2017; Barabás et al. 2018). Nevertheless, most of these challenges can be at least partially overcome by assessing multiple metrics and scales.

First, consider alternate stable states, which can confound tests if perturbations are sufficiently strong to overcome locally stable coexistence (Chesson 2000b; Levine et al. 2017; Saavedra et al. 2017; Barabás et al. 2018). In the PSF model, for example, invasion tests shift simulations into a state where both species can displace the other, but the two cannot coexist (fig. 2f; see sec. A2 of app. A; n.b.: although $r_0 > 0$, most ecologists would probably classify this result as unstable). In contrast, r_e indicates that species can coexist in the long term at most scales (fig. 3c). Jointly, these results (correctly) suggest that coexistence in the PSF model is locally stable but that coexistence is not robust to large disturbances.

Conversely, just as disturbances that are too large risk obscuring local stability, perturbations that are too small can be impossible to distinguish from background noise (May 1973; Turelli 1980). For example, r_e consistently predicts instability at small spatial extents ($s < 0.05$) regardless of underlying dynamics (fig. 3a–3e). This is because these extents harbor small populations, which allows demographic stochasticity to overwhelm effects of perturbations (see figs. S9, S10). Similar effects confound r_e at all spatial extents for the disturbance model, because species are frequently driven to low abundance (see sec. B2 of app. B and fig. S11). Results from invasion tests are less strongly influenced by stochasticity, as effects of larger perturbations are easier to detect (fig. 3f–3j). Thus, especially when populations are small, cases where both r_e and r_0 are positive likely indicate that long-term coexistence is possible but that dynamics are strongly influenced by stochastic fluctuations.

A related trade-off involves temporal extent. Effects of perturbations or invasions are typically diluted over time, making responses difficult to detect (Sheil and May 1996). A partial solution is to scale estimates of r_e and r_0 by t, as we do in figure 4, as these transformed metrics measure total recovery rather than recovery per unit time. If long-term estimates of $r_e t$ remain negative or of $r_0 t$ remain positive, then this indicates long-term coexistence. Tests conducted at shorter temporal extents also risk being confounded by transient or fluctuating dynamics, as is the case with r_0 for the RPS model (figs. 2h, 3i), or short temporal extents ($t < D$) in the disturbance model (fig. 3b, 3g). In general, we therefore advocate caution when interpreting results from systems that may be subject to transience or fluctuations if it is suspected that surveys are too short to accurately capture the full range of potential dynamics.

One last obstacle pertains to processes acting outside the scope of tests. For example, in the PSF and neutral models, r_0 incorrectly indicates stability at small scales because dispersal from outside the focal area overwhelms local dynamics (fig. 3i, 3j; Pimm 1984; Hubbell 2001). Similarly, r_e incorrectly indicates stability for the neutral model even at large spatial extents because total community abundance is stable (see fig. S12), which leads to compensatory increases in abundance that are shared across species (see fig. S13). While these effects could potentially be mitigated by preventing immigration into the focal patch or by replacing removed individuals with those of another species, such interventions would be difficult to implement in practice and could inadvertently destabilize the system, for example, by altering metacommunity processes (Tilman et al. 1994; Leibold and Chase 2018; see the example in fig. S14). In these cases, there is no general solution. Either statements about stability must be limited to the observed range of spatial and temporal extents or specific assumptions must be made about how the system behaves outside these extents (see "Identifying Mechanisms" below).

Identifying Mechanisms

In addition to providing qualitative information about system dynamics, fingerprints also appear to be useful for identifying potential underlying mechanisms that influence system dynamics. For example, our old field analyses suggest that the Levins-OF model better explains observed dynamics than does the neutral-OF model, in accordance with results from long-term studies at Cedar Creek (Gleeson and Tilman 1990; Tilman 1990, 1994; Clark et al. 2019). Likewise, despite superficial similarities among some fingerprints—for example, the Levins versus PSF model (fig. 3*a*, 3*c*)—we were typically able to distinguish among simulated models with high certainty, especially when considered across multiple extents (see app. D and fig. S15). Critically, once an appropriate model has been identified, its stability properties can be used as a proxy of those for the empirical system, which overcomes many of the problems discussed here (Ellner et al. 2019).

An important caveat is that all information that is contained in stability fingerprints is also available in time series data on species abundances. Thus, similar results could probably be achieved by fitting dynamic models to multiscale information about species abundances. However, doing so requires complex methods that are rarely applied in practice (Detto and Muller-Landau 2016; Clark et al. 2018). Fingerprints may therefore be useful as a low-dimensional summary of dynamic information, which is comparatively easy to interpret and compare across models. This feature could be particularly important in empirical systems where community composition is influenced by large-scale processes such as metacommunity dynamics, biogeography, or evolutionary history (Lawton 1999; Chesson 2000*b*; Davis and Shaw 2001; Ricklefs 2008; Leibold and Chase 2018). Ideally, information already collected as part of other studies may provide sufficient spatial and temporal extents to identify underlying mechanisms. However, for some systems or mechanisms, available data may not be sufficient to identify unique aspects of fingerprints, and surveying the necessary scales may not be feasible. Under these circumstances, fingerprinting may still help reduce the number of potential mechanisms that could explain observed patterns.

Future Directions

Here, we present an empirically tractable approach for characterizing long-term coexistence in ecological communities. By exploring some of the challenges that confront stability analysis, we hope that we have convinced readers that rigidly interpreted binary metrics of coexistence are probably not appropriate in most empirical contexts. More importantly, by demonstrating how fingerprinting can be applied to heuristically characterize system dynamics and stability, we hope

to encourage future empirical studies to employ a wider range of experimental perturbation treatments and to conduct measurements across a broader array of spatial and temporal extents.

The next challenge will be to find combinations of metrics, scales, and perturbation types that can uniquely identify a wide range of dynamic behavior across many systems (Levin 1992; Ives and Carpenter 2007). For some systems, this may require new measurements or methods for extrapolating dynamic behavior across unobserved spatial and temporal scales (Leibold and Chase 2018). However, in many cases it seems likely that the data, methods, and theory necessary for conducting these tests are already in place. We are therefore optimistic that stability fingerprinting could help greatly expand our understanding of coexistence in a rich variety of theoretical and empirical systems.

Data and Code Availability

All data and source code needed to reproduce the analyses and figures presented here are provided in a zip file (available online) in the archive "stability_scaling." Data used in this study can be accessed on the LTER network data portal (https://doi.org/10.6073/pasta/aa029df8f7a6091ea879ceb5c6673963; Knops 2018).

Literature Cited

Adler, P. B., A. Fajardo, A. R. Kleinhesselink, and N. J. B. Kraft. 2013. Trait-based tests of coexistence mechanisms. Ecology Letters 16:1294–1306.

Allesina, S., and J. M. Levine. 2011. A competitive network theory of species diversity. Proceedings of the National Academy of Sciences of the USA 108:5638–5642.

Anderson, H. M., V. Hutson, and R. Law. 1992. On the conditions for permanence of species in ecological communities. American Naturalist 139:663–668.

Armstrong, R. A., and R. McGehee. 1980. Competitive exclusion. American Naturalist 115:151–170.

Arnoldi, J.-F., A. Bideault, M. Loreau, and B. Haegeman. 2018. How ecosystems recover from pulse perturbations: a theory of short- to long-term responses. Journal of Theoretical Biology 436:79–92.

Arnoldi, J.-F., M. Loreau, and B. Haegeman. 2016. Resilience, reactivity and variability: a mathematical comparison of ecological stability measures. Journal of Theoretical Biology 389:47–59.

Barabás, G., R. D'Andrea, and S. M. Stump. 2018. Chesson's coexistence theory. Ecological Monographs 88:277–303.

Barabás, G., and A. Ostling. 2013. Community robustness in discrete-time periodic environments. Ecological Complexity 15:122–130.

Black, A. J., and A. J. McKane. 2012. Stochastic formulation of ecological models and their applications. Trends in Ecology and Evolution 27:337–345.

Carpenter, S., B. Walker, J. M. Anderies, and N. Abel. 2001. From metaphor to measurement: resilience of what to what? Ecosystems 4:765–781.

Chesson, P. 2000a. General theory of competitive coexistence in spatially-varying environments. Theoretical Population Biology 58:211–237.

———. 2000b. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics 31:343–366.

———. 2003. Quantifying and testing coexistence mechanisms arising from recruitment fluctuations. Theoretical Population Biology 64:345–357.

———. 2017. AEDT: a new concept for ecological dynamics in the ever-changing world. PLoS Biology 15:e2002634.

———. 2018. Updates on mechanisms of maintenance of species diversity. Journal of Ecology 106:1773–1794.

Chesson, P., and S. Ellner. 1989. Invasibility and stochastic boundedness in monotonic competition models. Journal of Mathematical Biology 27:117–138.

Chesson, P., and N. Huntly. 1997. The roles of harsh and fluctuating conditions in the dynamics of ecological communities. American Naturalist 150:519–553.

Chesson, P., and J. J. Kuang. 2008. The interaction between predation and competition. Nature 456:235–238.

Clark, A. T. 2017. Constraints and tradeoffs: toward a predictive, mechanism-based understanding of ecological communities. University of Minnesota, St. Paul.

Clark, A. T., M. Detto, H. C. Muller-Landau, S. A. Schnitzer, S. J. Wright, R. Condit, and S. P. Hubbell. 2018. Functional traits of tropical trees and lianas explain spatial structure across multiple scales. Journal of Ecology 106:795–806.

Clark, A. T., J. M. H. Knops, and D. Tilman. 2019. Contingent factors explain average divergence in functional composition over 88 years of old field succession. Journal of Ecology 107:545–558.

Davis, M. B., and R. G. Shaw. 2001. Range shifts and adaptive responses to Quaternary climate change. Science 292:673–679.

Detto, M., and H. C. Muller-Landau. 2016. Rates of formation and dissipation of clumping reveal lagged responses in tropical tree populations. Ecology 97:1170–1181.

Donohue, I., H. Hillebrand, J. M. Montoya, O. L. Petchey, S. L. Pimm, M. S. Fowler, K. Healy, et al. 2016. Navigating the complexity of ecological stability. Ecology Letters 19:1172–1185.

Donohue, I., O. L. Petchey, J. M. Montoya, A. L. Jackson, L. McNally, M. Viana, K. Healy, et al. 2013. On the dimensionality of ecological stability. Ecology Letters 16:421–429.

Durrett, R., and S. Levin. 1994. The importance of being discrete (and spatial). Theoretical Population Biology 46:363–394.

Ellner, S. P., R. E. Snyder, P. B. Adler, and G. Hooker. 2019. An expanded modern coexistence theory for empirical applications. Ecology Letters 22:3–18.

Gleeson, S. K., and D. Tilman. 1990. Allocation and the transient dynamics of succession on poor soils. Ecology 71:1144–1155.

Grilli, J., G. Barabás, M. J. Michalska-Smith, and S. Allesina. 2017. Higher-order interactions stabilize dynamics in competitive network models. Nature 548:210–213.

Grimm, V., and C. Wissel. 1997. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. Oecologia 109:323–334.

Hillebrand, H., S. Langenheder, K. Lebret, E. Lindström, Ö. Östman, and M. Striebel. 2018. Decomposing multiple dimensions of stability in global change experiments. Ecology Letters 21:21–30.

Hubbell, S. 2001. The unified neutral theory of biodiversity and biogeography. Princeton Monographs in Population Biology. Princeton University Press, Princeton, NJ.

Inouye, R. S., N. J. Huntly, D. Tilman, J. R. Tester, M. Stillwell, and K. C. Zinnel. 1987. Old-field succession on a Minnesota sand plain. Ecology 68:12–26.

Ives, A. R., and S. R. Carpenter. 2007. Stability and diversity of ecosystems. Science 317:58–62.

Knops, J. 2018. Plant species percent cover data: successional dynamics on a resampled chronosequence. Environmental Data Initiative. https://doi.org/10.6073/pasta/aa029df8f7a6091ea879ceb5c6673963.

Lawton, J. H. 1999. Are there general laws in ecology? Oikos 84:177.

Lehman, C., A. Keen, and R. Barnes. 2012. Trading space for time: constant-speed algorithms for managing future events in scientific simulations. Page 1 *in* Proceedings of the International Conference on Scientific Computing (CSC). Steering Committee of the World Congress in Computer Science, Computer Engineering, and Applied Computing (WorldComp).

Leibold, M. A., and J. M. Chase. 2018. Metacommunity ecology. Vol. 59. Princeton University Press, Princeton, NJ.

Levin, S. A. 1992. The problem of pattern and scale in ecology: the Robert H. MacArthur Award Lecture. Ecology 73:1943.

Levine, J. M., J. Bascompte, P. B. Adler, and S. Allesina. 2017. Beyond pairwise mechanisms of species coexistence in complex communities. Nature 546:56–64.

Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. Bulletin of the Entomological Society of America 15:237–240.

MacArthur, R. H. 1970. Species packing and competitive equilibrium for many species. Theoretical Population Biology 1:1–11.

May, R. M. 1973. Stability and complexity in model ecosystems. Monographs in Population Biology 6. Princeton University Press, Princeton, NJ.

———. 1974. Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos. Science 186:645–647.

Meyer, K., A. Hoyer-Leitzel, S. Iams, I. Klasky, V. Lee, S. Ligtenberg, E. Bussmann, et al. 2018. Quantifying resilience to recurrent ecosystem disturbances using flow-kick dynamics. Nature Sustainability 1:671–678.

Murdoch, W. W. 1994. Population regulation in theory and practice. Ecology 75:271–287.

Pimm, S. L. 1984. The complexity and stability of ecosystems. Nature 307:321–326.

Ricklefs, R. E. 2008. Disintegration of the ecological community. American Naturalist 172:741–750.

Saavedra, S., R. P. Rohr, J. Bascompte, O. Godoy, N. J. B. Kraft, and J. M. Levine. 2017. A structural approach for understanding multispecies coexistence. Ecological Monographs 87:470–486.

Sheil, D., and R. M. May. 1996. Mortality and recruitment rate evaluations in heterogeneous tropical forests. Journal of Ecology 84:91–100.

Siepielski, A. M., and M. A. McPeek. 2010. On the evidence for species coexistence: a critique of the coexistence program. Ecology 91:3153–3164.

Stommel, H. 1963. Varieties of oceanographic experience: the ocean can be investigated as a hydrodynamical phenomenon as well as explored geographically. Science 139:572–576.

Suding, K. N., W. Stanley Harpole, T. Fukami, A. Kulmatiski, A. S. MacDougall, C. Stein, and W. H. van der Putten. 2013. Consequences of plant-soil feedbacks in invasion. Journal of Ecology 101:298–308.

Tilman, D. 1990. Constraints and tradeoffs—toward a predictive theory of competition and succession. Oikos 58:3–15.

———. 1994. Competition and biodiversity in spatially structured habitats. Ecology 75:2–16.

———. 2004. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. Proceedings of the National Academy of Sciences of the USA 101:10854–10861.

Tilman, D., R. May, C. Lehman, and M. Nowak. 1994. Habitat destruction and the extinction debt. Nature 371:65–66.

Turelli, M. 1980. Niche overlap and invasion of competitors in random environments II. The effects of demographic stochasticity. Pages 119–129 *in* W. Jäger, H. Rost, and P. Tautu, eds. Biological growth and spread. Vol. 38. Springer, Berlin.

———. 1981. Niche overlap and invasion of competitors in random environments I. Models without demographic stochasticity. Theoretical Population Biology 20:1–56.

———. 1986. Stochastic community theory: a partially guided tour. Pages 321–339 *in* T. G. Hallam and S. A. Levin, eds. Mathematical ecology. Vol. 17. Springer, Berlin.

Zelnik, Y. R., J.-F. Arnoldi, and M. Loreau. 2018. The impact of spatial and temporal dimensions of disturbances on ecosystem stability. Frontiers in Ecology and Evolution 6:224.

References Cited Only in the Online Enhancements

Anderson, G. W., A. Guionnet, and O. Zeitouni. 2010. An introduction to random matrices. Cambridge Studies in Advanced Mathematics. Cambridge University Press, New York.

Bever, J. D., K. M. Westover, and J. Antonovics. 1997. Incorporating the soil community into plant population dynamics: the utility of the feedback approach. Journal of Ecology 85:561–573.

Borer, E. T., W. S. Harpole, P. B. Adler, E. M. Lind, J. L. Orrock, E. W. Seabloom, and M. D. Smith. 2014. Finding generality in ecology: a model for globally distributed experiments. Methods in Ecology and Evolution 5:65–73.

Chesson, P. 1990. MacArthur's consumer-resource model. Theoretical Population Biology 37:26–38.

Clark, A. T., and C. Neuhauser. 2018. Harnessing uncertainty to approximate mechanistic models of interspecific interactions. Theoretical Population Biology 123:35–44.

Cushing, E. J. 1963. Late-Wisconsin pollen stratigraphy in east-central Minnesota. University of Minnesota, Minneapolis.

Dormann, C. F. 2008. On community matrix theory in experimental plant ecology. Web Ecology 8:108–115.

Kitajima, K., and D. Tilman. 1996. Seed banks and seedling establishment on an experimental productivity gradient. Oikos 76:381–391.

Muller-Landau, H. C. 2010. The tolerance-fecundity trade-off and the maintenance of diversity in seed size. Proceedings of the National Academy of Sciences of the USA 107:4242–4247.

R Development Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Tilman, D. 1987. Secondary succession and the pattern of plant dominance along experimental nitrogen gradients. Ecological Monographs 57:189–214.

Tilman, D., J. Knops, D. Wedin, P. Reich, M. Ritchie, and E. Siemann. 1997. The influence of functional diversity and composition on ecosystem processes. Science 277:1300–1302.

Weisser, W. W., C. Roscher, S. T. Meyer, A. Ebeling, G. Luo, E. Allan, H. Beßler, et al. 2017. Biodiversity effects on ecosystem functioning in a 15-year grassland experiment: patterns, mechanisms, and open questions. Basic and Applied Ecology 23:1–73.