Scaling-up biodiversity-ecosystem functioning research

Authors: Andrew Gonzalez1*, Rachel M. Germain2, Diane S. Srivastava3, Elise Filotas, Laura E. Dee, Dominique Gravel5, Patrick L. Thompson2, Forest Isbell6, Shaopeng Wang7, Sonia Kéfi8, Jose Montoya9, Yuval R. Zelnik9, Michel Loreau9

1Dept. Biology, McGill University, 1205 Dr. Penfield Avenue, Montreal, H3A 1B1, Canada.

2Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, Canada, V6T 1Z4.

3Center for Forest Research, Département Science et Technologie, Université du Québec, Téluq, 5800 rue Saint-Denis, Montréal, Québec H2S 3L5, Canada.

4Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado 80309 USA.

5Département de biologie, Université de Sherbrooke, 2500 Boul. De l’Université, Sherbrooke, J1K2R1, Canada.

6Department of Ecology, Evolution, and Behavior, University of Minnesota, 1479 Gortner Avenue, St. Paul, MN 55108, USA.

7Institute of Ecology, College of Urban and Environmental Science, and Key Laboratory for Earth Surface Processes of the Ministry of Education, Peking University, 100871 Beijing, China

8ISEM, CNRS, Univ. Montpellier, IRD, EPHE, Montpellier, France.
Centre for Biodiversity Theory and Modelling, Theoretical and Experimental Ecology Station, CNRS, 2 route du CNRS, 09200 Moulin, France.

Author emails: andrew.gonzalez@mcgill.ca; rgermain@zoology.ubc.ca; srivast@zoology.ubc.ca; elise.filotas@teluq.ca; Laura.De@colorado.edu; dominique.gravel@usherbrooke.ca; patrick.thompson@zoology.ubc.ca; jcowles@umn.edu; isbell@umn.edu; sonia.kefi@umontpellier.fr; josemaria.montoyateran@sete.cnrs.fr; yuval.zelnik@sete.cnrs.fr; michel.loreau@sete.cnrs.fr

Running title: Scaling-up biodiversity-ecosystem function

Keywords: scale, ecosystems, biological diversity, ecosystem functioning, beta diversity, environmental heterogeneity, turnover

Type of Article: Reviews & Synthesis

Words in abstract (202); main text (7687); number of references (207); number of figures (5); number of tables (1); number of text boxes (1; 644 words).

* Corresponding author: Andrew Gonzalez 1205 Doctor Penfield, Department of Biology, McGill University, Montreal QC, H3A 1B1, Canada; phone 1-514-398-6444; fax 1-514-398-5069.

Statement of authorship: All authors contributed to the idea and design of the project. AG lead the writing. EF produced the data and design for Figure 5. All authors contributed to the writing of the manuscript.

Data accessibility statement: The data used to produce Figure 5 will be available online in open access, and archived on an online repository upon manuscript acceptance.
Abstract

A rich body of knowledge links biodiversity to ecosystem functioning (BEF), but it is primarily focused on small scales. We review current theory and identify six expectations for scale dependence in the BEF relationship: 1) a nonlinear change in the slope of the BEF relationship with spatial scale; 2) a scale-dependent relationship between ecosystem stability and spatial extent; 3) coexistence within and among sites will result in a positive BEF relationship at larger scales; 4) temporal autocorrelation in environmental variability affects species turnover and thus the change in BEF slope with scale; 5) connectivity in metacommunities generates nonlinear BEF and stability relationships by affecting the synchrony in dynamics at local and regional scales; 6) spatial scaling in food web structure and diversity will generate scale dependence in ecosystem functioning. We suggest directions for synthesis that combine approaches in metaecosystem and metacommunity ecology and integrate cross-scale feedbacks. Tests of this theory may combine remote sensing with a generation of networked experiments that assess effects at multiple scales. We also show how anthropogenic land cover change may alter the scaling of the BEF relationship. New research on the role of scale in BEF will guide policy linking the goals of managing biodiversity and ecosystems.
1. INTRODUCTION

Life has transformed the Earth, mediating fluxes of elements and energy from the smallest to the largest spatial scales (Schramski et al., 2015; Le Quéré et al., 2016). The diversity and distributions of plant, animal, and microbial life reflect evolutionary and ecological processes constrained by broad-scale abiotic gradients of energy, resources and meteorological conditions on land (Hawkins et al. 2003; Kreft & Jetz 2007; Pappas et al. 2017) and in the oceans (Vallina et al. 2014; Woolley et al. 2016; Frainer et al. 2017; Tréguer et al. 2018). Even while the distribution of biodiversity reflects gradients of energy and limiting resources, it also contributes to how effectively those gradients are exploited to confer ecosystem functioning, such as variability in the rates of primary and secondary production (Baldocchi 2014; Niu et al. 2017; Pappas et al. 2017; Jia et al. 2018).

Yet, understanding how feedbacks between biodiversity and ecosystem functioning occur, and vary from local to biogeographic scales, is a major challenge (Enquist et al. 2003, 2007; Grace et al. 2007; Gross & Cardinale 2007; Violle et al. 2014; Guidi et al. 2016; Maestre et al. 2016; Tréguer et al. 2018; Bagousse-Pinguet et al. 2019), one that is urgent to resolve as biodiversity change occurs at multiple scales in response to climate warming, species introductions, and habitat degradation (Reichstein et al. 2014; Snelgrove et al. 2014; Isbell et al. 2017; Chase et al. 2019).

Biodiversity-ecosystem functioning (BEF) has focused on isolating the causal pathways by which biodiversity change alters the magnitude and stability of ecosystem processes (Hooper et al. 2005; Cardinale et al. 2011). Theory has played a major role in establishing predictions and validating interpretations of data. An important example of this is the way BEF effects can arise from selection versus complementarity effects at local scales (Loreau et al. 2001; Tilman, 1997). Although much of the focus in this area over the last three decades has built upon theory for fine-
scale and short time periods in ecological systems, there are theoretical expectations for how
spatial and temporal niche complementarity vary in importance over time, and increase in
importance at greater scales on land and in the oceans (Cermeño et al. 2016; Vallina et al. 2017).

Meta-analyses of hundreds of BEF experiments have shown consistent relationships between
biodiversity and ecosystem functioning across different ecosystem types and functions (Balvanera
et al. 2006; Cardinale et al. 2011; O’Connor et al. 2017). However, this body of work has
emphasized a limited range of spatial and temporal scales; experimental plots cover an area of ~1-
100m$^2$ and have lasted ~1-10 generations (see Fig. 1, see also Cardinale et al., 2009). This means
that while these studies can test the underlying mechanisms of BEF and short-term predictions of
the theory, they cannot also directly address theoretical predictions that extend BEF relationships
to broader scales (Loreau et al. 2003; Isbell et al. 2017, 2018; Thompson et al. 2018). A new
generation of studies is starting to provide a deeper understanding of BEF at larger scales, in more
realistic settings, across ecosystem types and gradients of climate (Duffy et al. 2007, 2017; Fung
et al. 2015; Barnes et al. 2016; Liang et al. 2016; Oehri et al. 2017; Ratcliffe et al. 2017; Winfree
et al. 2018; Bagousse-Pinguet et al. 2019; DeLong & Gibert 2019; Lefcheck et al. 2019), but a
systematic assessment of theoretical predictions to bolster the interpretation of this new generation
of empirical studies is lacking.

We are left, therefore, with one of the challenges associated with BEF research still unresolved: to
what extent does the strength of the relationship linking change in biodiversity to change in
ecosystem functioning depend on scale (Bengtsson et al. 2002; Naeem 2006)? How can existing
terapy be used to scale-up our understanding of the BEF relationship obtained at small spatial
scales (Loreau et al. 2003; Cardinale et al. 2004; Burley et al. 2016; Yasuhara et al. 2016; Isbell et al. 2018; Thompson et al. 2018)? These questions, which we label ‘the question of scale for BEF’ is the focus of this paper. Articulated differently, do the processes explaining predominantly positive BEF correlations at local scales extend to regional and global scales (Ptacnik et al. 2008; Snelgrove et al. 2014; Vallina et al. 2014; Violle et al. 2014; Liang et al. 2016; Wang et al. 2017; Acevedo-Trejos et al. 2018; Delsol et al. 2018; García-Palacios et al. 2018) and decadal and centennial time scales, or do different processes dominate at different scales (Brose & Hillebrand 2016; Isbell et al. 2017)? Here we explore how answers to these questions can be sought with existing and new theory, and subsequently tested by experiments and multiscale observations.

In this article, we review and synthesize disparate theories addressing how the BEF relationship varies with spatial, temporal and organizational scale (Fig. 1). Box 1 provides an overview of the most relevant aspects of scale. We highlight expectations from scaling theory that address the relationships between biodiversity, ecosystem functioning and stability, and process-based theories of BEF that formalize causal relationships. These theories involve different assumptions, and so provide complementary explanations for why biodiversity and ecosystem functioning should vary across scales. We assess the extent to which these theoretical predictions have been supported by empirical observations and experiments. Our review of the empirical literature is not exhaustive, but highlights the evidence addressing this theory. We do not restrict our review to particular measures of biodiversity or ecosystem functioning: the theory we review encompasses expectations for measures of richness and diversity of species (or functional groups) and their interactions across levels of organization, and relates them to measures of functioning that include both ecosystems stocks (e.g. biomass) and processes. We provide several avenues for theoretical
and empirical synthesis. From here, we discuss how humans may be altering biodiversity and ecosystem function at different scales and provide a detailed example for landcover change. We close with recommendations for testing the theory with new datasets derived from molecular methods, networked experiments and remote sensing.

2. EXPECTATIONS FROM THEORY SCALING BEF AND STABILITY

A first task is to assess how the biodiversity, ecosystem functioning, and stability depend on scale without evoking the ecological processes that generate them. One reason that directly ‘scaling up’ patterns from fine scale theory and evidence may not be possible is the observation that area (or volume, in aquatic environments) influences the relative abundance of individuals and species, the ways species diversity and biomass accumulate in space, and the correlations in their biomass fluctuations in space and time. In this view, any variation in BEF relationships across scales could arise from the scale dependence in the distributions of individuals within and among species. Two such approaches to scale dependence (see Box 1 for important definitions) have received attention: the first considers a decomposition of space into two scales—local and regional—and describes turnover in local assemblage biodiversity and ecosystem function as observations are aggregated from local sites into coarser grains. The second treats space as a continuum along a transect or expanding spatial extent (see Fig. 2; Barry et al. 2019). In the following subsections we summarize the key findings from each approach.

2.1 Expectation 1: The slope of the BEF relationship is scale dependent

We consider whether the nonlinear BEF relationship commonly observed at small spatial grains changes as we aggregate observations to encompass larger grains (Fig. 2, Fig. 3). Specifically, as
we aggregate small spatial units of area (grain) to form larger spatial units, we consider how estimates of the form (slope) and explanatory power (e.g. $R^2$) of the BEF relationship change up to the largest spatial unit. A useful starting point here is to consider when the BEF relationship might remain the same at large scales as it is at small scales. Thompson et al. (2018), building on Cardinale et al. (2004), explored this question and found that the BEF relationship should remain constant with scale only if a proportional change in biodiversity results in the same proportional change in ecosystem functioning at all scales. This requires that three unlikely conditions are met: 1) local ($\alpha$) species richness is constant across the entire region, 2) the local ($\alpha$) scale slope of the BEF relationship is constant across the entire region, and 3) there is either complete overlap in composition across all local communities or no compositional overlap between local communities.

When conditions 1 and 2 are not met, scale dependence of the BEF slope arises from non-linear averaging of spatially heterogeneous values of species richness (condition 1) or EF (condition 2). This, however, has a relatively modest effect on scaling of the BEF relationship compared to violation of condition 3, compositional turnover across space where changes in mean $\alpha$ richness do not result in the same proportional change in $\gamma$ richness. Consequently, when fine-scale variation in the BEF relationship is aggregated, the change in BEF slope becomes nonlinear (Fig. 3A). One untested theoretical expectation is that similar patterns may arise when aggregation is done through time instead of space. This expectation arises because both the species-time relationship, and the species time-area relationship, show temporal turnover in slopes that are very similar in form to the species-area relationship (Adler & Lauenroth 2003).

Empirical evidence
Most BEF experiments or surveys use a single small plot size, which can reveal large geographic plot-to-plot variation in the slope of BEF relationship (Liang et al., 2016). However, a growing number of studies have assessed how the slope of the BEF relationship changes when at least two spatial grains are examined (Roscher et al. 2005; Costanza et al. 2007; Chisholm et al. 2013; Sullivan et al. 2017; Sanaei et al. 2018; Luo et al. 2019). For example, Chisholm et al. (2013) studied the effect of tree species richness on forest biomass and productivity in 25 forest plots varying in areal extent from 8 to 50 ha. They systematically varied spatial grain (0.04 to 1 ha) and found that the relationship between tree species richness, above-ground biomass, and coarse woody dry productivity changed qualitatively with grain. Species richness and niche complementarity effects were dominant predictors of ecosystem properties at small spatial grains, while environmental gradients explained variation at larger grains. At the smallest grain, 0.04 ha, doubling species richness corresponded to a 48% increase on average in productivity and a 53% increase in above-ground biomass. But at larger spatial grains (0.25 and 1 ha) the average BEF relationship was only weakly positive (doubling diversity led to a 5% and 7% increase in productivity and biomass, respectively), and in fact negative relationships were more common. Biomass and productivity were positively correlated across spatial grains. Sullivan et al. (2017) also conducted a multi-scale evaluation of diversity-carbon relationships in tropical forests across the tropics in three continents. Diversity-carbon relationships among all plots at 1 ha scale were absent, and within continents were either weak (Asia) or absent (Amazonia, Africa).

Testing theoretical expectations for BEF with empirical studies is challenging because of the covariance of other factors, such as climate or productivity, that can mask changes in the BEF relationship (Loreau 1998). This is especially true when accounting for variation in the BEF slope.
at increasingly larger scales. For example, Costanza et al. (2007) found that the correlation between vascular plant richness and net primary production (NPP) at two scales – site and ecoregion – in North America depends on climate. At the site scale, 57% of the variance in NPP was correlated with variation in richness after accounting for the effects of climate. In contrast, at the ecoregion scale, the BEF relationship was found to change sign over three ranges of temperature (negative at low temperatures (-2°C average), no correlation at mid-temperatures (5°C average), and positive at high temperatures (13°C average). Without species composition data it is difficult to assess whether this result occurred via condition 3 identified by Thompson et al. (2018).

2.2 Expectation 2: Stability of function scales nonlinearly with area

Observed BEF relationships may vary with scale if the stability of this relationship varies, even while the underlying BEF relationship does not vary. One way to quantify stability is as invariability – that is, low temporal variation in population or community biomass. Invariability, like many other properties of ecosystems (most notably, species richness), increases with area (spatial extent). The key expectation from work on the invariability-area relationship (IAR) is that aggregate biomass and its variability scale nonlinearly with area (Wang et al., 2017). Wang et al. (2017) found that, like the species-area relationship (SAR), the IAR can have a triphasic form (Fig. 3B), the shape of which is related to the SAR (Delsol et al. 2018).

The key finding from work on the IAR is that the scaling of ecosystem stability with area is governed by the spatial asynchrony in species’ biomass fluctuations. Asynchrony in biomass fluctuations, together with the spatial distribution of individuals and species (sub panels in Fig 3B), determine the shape of IARs (Fig. 3B main plot). These two facets of the IAR describe how
quickly spatial averaging of temporal variability occurs with increasing area. In the limiting case where the biomass of individuals fluctuates synchronously within species, but independently among species, the IAR coincides exactly with the SAR because species’ identity governs the changes in both the number of species and invariability. In other theoretical scenarios, where the synchrony in biomass fluctuations within species are assumed to decay with distance, IARs become disconnected from SAR (Delsol et al. 2018).

Although work on the IAR has focused on temporal variation of biomass production so far, we also expect spatial invariability to show a positive IAR. Increased invariability with grain size should ultimately result from the fact that the mean level of EF per area stays constant while its standard deviation is expected to decrease (Chave et al. 2004). Asynchronously fluctuating variables compensate for each other at larger scales (Loreau 2010), and this should hold in both space and time.

**Empirical evidence**

Two large-scale datasets have been examined for triphasic IARs (Wang et al. 2017). Global primary productivity (MODIS data) across five continents exhibits triphasic curves, characterized by steeper increases in invariability at both small and large scales as predicted (Wang et al. 2017). However, this observation was not found in the North American breeding bird survey, possibly because the bird survey was a partial assessment at a subcontinental extent. More datasets are needed to assess whether the IAR is consistently triphasic. An analysis of IAR across trophic levels could be achieved with the bird survey or marine food web data (McGinty et al. 2012).
3. PROCESS-BASED EXPECTATIONS FOR SCALE-DEPENDENCE IN BEF IN SPACE AND TIME

So far, we have described general expectations from theory that emphasizes the statistical properties of diversity, fluctuations and function across scale, without considering scale dependent ecological processes. These include theory for how species interactions, temporal and spatial environmental variability, and metacommunity processes affect BEF at different scales. A review of the literature revealed a number of relevant ecological processes, which we summarize in Table 1. We now address four of these expectations in greater depth and discuss empirical studies providing support for them.

3.1 Expectation 3: Coexistence within and among sites will result in a positive BEF relationship at larger scales

A general expectation from coexistence theory is that larger scales of space and time encompass a greater range of environments, increasing species’ opportunities for niche partitioning and therefore BEF relationships that extend to larger scales. Early work used resource competition theory to articulate when species complementarity due to niche differentiation will explain overyielding in plant communities at small scales (Tilman et al. 1997). However, coexistence mechanisms are inherently scale dependent (Hart et al. 2017). If we assume species differ in average fitness in different environments and that no species can exist in all environments, environmental heterogeneity is expected to promote ecosystem functioning across space via spatial niche complementarity (Williams et al. 2017) and through time via temporal and spatio-temporal niche complementarity (Chesson et al. 2001; Gross & Cardinale 2007). Even in the absence of coexistence at local scales, spatial variation in species dominance (e.g. Winfree et al., 2018) can
result in niche complementarity at large spatial scales, and therefore generate a positive BEF relationship at those scales. The fact that biodiversity can underpin ecosystem functioning via complementarity within and among environments means that BEF effects are likely important at large scales.

Although biodiversity might increase ecosystem functions in each of a given set of environments (Tilman et al. 2012), its effects are ultimately constrained by limiting resources and physiological constraints that those environments impose (Harpole & Tilman 2007). If considered relative to fluctuating environmental conditions, the effect of biodiversity on some ecosystem functions, such as primary production, might be weak and difficult to isolate. However, larger spatial scales encompass a greater range of microgeographic heterogeneity (e.g., soil depth), habitat types (e.g., grassland vs. forest), and climates (Bell & Lechowicz 1991). Thus, if considered relative to an increase in biodiversity at small scales, the greater range of environments encompassed at larger scales adds more opportunity for niche partitioning (Ritchie & Olff 1999; Leibold & Chase 2018) which should strengthen BEF relationships.

**Empirical evidence**

Experiments that directly manipulate coexistence via environmental heterogeneity and examine effects on BEF relationships at different scales are scarce (Langenheder et al. 2010; Gravel et al. 2011). Gravel et al. (2011a) evolved assemblages of generalist and specialist marine bacteria and assessed their ability to metabolize a range of carbon substrates. They found that assemblages of generalists were more productive on average because of their superior ability to exploit the imposed heterogeneity in the resource environment. However, the slope of the BEF relationship
was stronger for the assemblages of specialists because of enhanced niche complementarity. A number of experiments have manipulated habitat heterogeneity and examined effects on BEF relationships (Tylianakis et al. 2008). Experiments with naturally-occurring species pools demonstrate a strengthening of the effect of biodiversity on ecosystem function with habitat heterogeneity, for example, the effect of algal diversity on stream water quality strengthens with substrate complexity (Cardinale 2011). Similarly, in the rocky intertidal, algal grazer species differentially feed in the presence and absence of barnacles, leading to increased spatial complementarity of algal consumption when barnacles are patchily distributed (Whalen et al. 2016).

3.2 Expectation 4: Autocorrelation in the environment will slow the rate of saturation in BEF relationship

Environmental variability on land and in the oceans is characterized by its autocorrelation (Steele 1985). Autocorrelation (i.e., the slow decay in environmental similarity in time or space) influences many ecological properties (Vasseur & Yodzis 2004), including rates of species turnover (β diversity) and productivity (Storch et al. 2007). As we saw in section 2, species turnover is a key factor governing scale dependence in the strength of the BEF relationship. A follow-on expectation is that the degree of scale dependence in BEF is mediated by environmental autocorrelation. If species replace each other over time and space in response to environmental fluctuations, then autocorrelation sets the rate of species turnover, which in turn sets the scale over which the BEF relationship saturates.
This prediction that environmental autocorrelation sets the scale over which the BEF relationship saturates can be evaluated using simulations of species competing for resources exposed to stochastic environmental fluctuations. Gonzalez and De Feo (2007) tested this prediction and found that the magnitude and stability of functioning in competitive communities depend strongly upon the degree of temporal environmental autocorrelation (Fig. 4). In the absence of temporal autocorrelation (white noise), community dynamics were characterized by high frequency, small-amplitude population fluctuations, and biomass was evenly distributed across species over all time scales (Fig. 4A). With increasing autocorrelation (Fig. 4B, C), the environment changed state more slowly, driving population dynamics with periods of alternating dominance and low evenness over short time scales. Increasing species richness increased biomass production (Fig. 4D) and stability (Fig. 4E) in all environment types, but – as predicted above – the effect of diversity was most important under autocorrelated conditions (Gonzalez & De Feo 2007).

A second finding of Gonzalez & De Feo’s (2007) model is that the scale of environmental autocorrelation will determine how many species are needed to reach a given level of EF. The slow turnover in diversity in autocorrelated environments should result in a slower saturation of the BEF slope compared to uncorrelated environments, where the full range of environmental variance is experienced over short intervals of space and time. Because of the near ubiquity of autocorrelated environmental conditions across scales (Steele 1985; Bell et al. 1993; Vasseur & Yodzis 2004), we expect it to be a strong determinant of scale-dependence in BEF processes.

Empirical evidence
Very little empirical evidence exists for spatial or temporal autocorrelation as a determinant of scale-dependence in BEF. Using plankton microcosms, Descamps-Julien and Gonzalez (2004) showed that autocorrelated fluctuations in temperature had a greater stabilizing effect on community algal biomass than uncorrelated fluctuations. This occurred because autocorrelated temperature variation allowed different species to more easily track the changing temperatures and these differential species’ responses result in lowered covariance in total community biomass. This theory could be tested by analyses of variation in freshwater and marine plankton diversity where spatial data and time series of primary production and physical environmental parameters mediating diversity are available, and could be used to estimate the variance spectra of these processes (Lévy 2008; Lévy Marina et al. 2015; Smith et al. 2016; Soininen et al. 2016; Tréguer et al. 2018).

3.3 Expectation 5: Connectivity has nonlinear effects on the strength and stability of BEF across scales

Spatial models predict BEF effects at multiple scales and that feedbacks across scales arise because of the transfer of organisms and resources among ecosystems (Peterson et al. 1998; Peters et al. 2007). Metacommunity and metaecosystem theories exemplify these feedbacks by showing that the direction and rates of dispersal govern local and regional biodiversity, and the rate and stability of biomass production and resource use (Loreau et al. 2003; Marleau et al. 2014; Thompson and Gonzalez, 2016; Leibold et al., 2017; Thompson et al., 2017). The movement of individuals and resources in these models causally links BEF across scales but is rarely studied in BEF experiments.
The spatial insurance hypothesis (SIH; Loreau et al., 2003; Shanafelt et al., 2015; Thompson and Gonzalez, 2016) states that dispersal links coexistence to the magnitude and stability of single and multiple ecosystem functions. The key prediction of the SIH is that changing connectivity (e.g. dispersal rate) leads to a nonlinear (e.g., unimodal) effect on functioning and stability. At very low connectivity species cannot move from patch-to-patch to effectively track their environmental optima, and so diversity is lost (i.e. due to local competitive exclusion). Intermediate rates of movement promote species persistence because they can track shifting environmental conditions. This spatial sorting of species results in turnover in species dominance which enhances biomass production when the environment is fluctuating locally and regionally (Thompson et al., 2017). Intermediate rates of movement also maintain local diversity by mass effects. In many cases, although this is not universal (Haegeman & Loreau 2014), at very high rates of dispersal a few species come to dominate the region because of competitive exclusion by species that have the greatest fitness for average conditions across all patches. Intermediate rates of dispersal therefore strongly stabilize productivity and resource use because of asynchronous species fluctuations. Spatial insurance theory generalizes the local temporal insurance that occurs within patches (Yachi & Loreau 1999). The insurance results from differences among species in their responses to fluctuating environmental conditions (Elmqvist et al. 2003). Both temporal and spatial insurance provide stabilizing effects to regional ecosystems and reflect the effects of $\alpha$ (local) and $\beta$ (among community) diversity on ecosystem properties. Specifically, while $\alpha$ diversity decreases local ecosystem variability, $\beta$ diversity generally contributes to increasing spatial asynchrony among local ecosystems, as shown by Wang and Loreau (2016) using Lotka-Volterra multi-patch metacommunity models. In an important link to expectation 1 and 4, such metacommunity models
simultaneously show that, at the regional scale, the stabilising effect of $\beta$ diversity increases as spatial environmental autocorrelation increases (Wang and Loreau 2016).

Empirical evidence

Several studies have experimentally controlled connectivity and shown that it affects diversity, EF and stability (France & Duffy 2006; Staddon et al. 2010; Haddad et al. 2015; Thompson et al. 2015; Guelzow et al. 2017; Limberger et al. 2019). One BEF experiment that explicitly assessed scale while controlling dispersal in a metacommunity. Venail et al. (2010) assembled a metacommunity composed of a number of genotypes of bacterium *Pseudomonas fluorescens* to examine BEF relationships under varying dispersal rates in a spatially heterogeneous landscape. A BEF relationship was only observed at the regional scale—the scale at which resource heterogeneity allowed spatial complementarity. Spatial complementarity peaked at intermediate dispersal rates, the rate that allowed species to access and persist in all suitable local environments. At local scales, a single carbon source precluded niche differentiation so BEF relationships did not emerge. Dispersal increased diversity through mass effects, but not productivity because local resource complementarity was not possible. Experiments beyond the lab are needed and could be done in grasslands where dispersal can be quantified and spatial plot configuration controlled to reveal the effects on BEF (Germain et al. 2017). These results demonstrate how the scales at which BEF relationships emerge depend on the scales of interaction between ecological processes; in this case, dispersal, competition, and environmental sorting.

3.4 Expectation 6: Interaction network structure will influence scale dependence of EF in food webs
So far, our discussion of the scaling of BEF relationships accounts for biodiversity change at a single trophic level, ignoring the additional complexity that emerges from network measures of diversity. Most theoretical and empirical investigations of BEF in food webs have focused on small spatial scales (Duffy 2003; Thébault & Loreau 2003; Loreau & Holt 2004), limiting our ability to assess how spatial scale affects ecosystem functioning in meta-networks (Barnes et al. 2016). However, progress has been made on three fronts: 1) interaction network diversity and dissimilarity in space and time (Brose et al. 2004; Poisot et al. 2013; Barnes et al. 2016; Schieber et al. 2017; Galiana et al. 2018), 2) causal relationships between food web structure, diversity and ecosystem functions (Poisot et al. 2013; Wang & Brose 2018), and 3) the ecosystem effects of trophic coupling by mobile consumers (McCann et al. 2005; Marleau & Guichard 2014).

There are several reasons why network diversity and structure vary with spatial scale. First, food chain length is expected to increase with habitat area or volume (Holt et al. 1999; Post et al. 2000), leading to different SARs at each trophic level (Holt et al. 1999; Ryberg & Chase 2007). Moreover, generalist species do better than specialist species on small and less connected areas because they are more likely to meet their energy requirements than specialist species. These outcomes lead to a network connectance-area relationship (Gravel et al. 2011). Pillai et al. (2011) also found that the complexity of food web topology, in particular the prevalence of omnivory and intraguild predation, should increase with sampled area. This complexity can in turn increase species diversity and ecosystem functioning and also strengthen BEF relationship in food webs (Wang et al. 2019).
Recent theory suggests that network properties can be related to BEF mechanisms. These include two ideas: trophic complementarity (Poisot et al. 2013) and the vertical diversity (Wang and Brose, 2018). Trophic complementarity extends the species resource complementarity concept to a trophic network by indicating how much consumers in a network feed on different prey species. Maximal complementarity occurs in trophic networks when consumers have low overlap in resource use and predators have low overlap in their exploitation of consumers, e.g. a food web made entirely of unconnected linear food chains (Poisot et al. 2013).

The vertical diversity hypothesis (Wang and Brose 2018) predicts that, at a given level of nutrient supply, primary production increases with vertical diversity of complex food webs, as measured by the trophic level and/or body size of the top predator. The vertical diversity hypothesis is explained by the top–down regulation imposed by the vertical diversity on plant species, which induces selection and complementarity effects analogous to those of horizontal diversity in single-trophic BEF studies. To date, little theory has directly addressed the scaling of food web structure and measures of ecosystem functioning, whether for total energy flux (Brose et al. 2004; Barnes et al. 2014; Wang & Brose 2018), or stability (McCann et al. 2005; Marleau & Guichard 2014).

**Empirical evidence**

Evidence is accruing that the manipulation of horizontal and vertical network diversity affects stability and ecosystem function (Srivastava & Bell 2009; Thibaut et al. 2012; Fornoff et al. 2019; Zhao et al. 2019), but few studies have incorporated scale. Experiments have found that the magnitude and stability of BEF is modified by changes in food web diversity and spatial scale (France & Duffy 2006; Staddon et al. 2010; Limberger et al. 2019). These studies connected
habitat patches via dispersal to study the emergent relationships between B and EF across patches in a metacommunity. France and Duffy (2006) found that dispersal and grazer diversity temporally destabilized the biomass of primary producers in local patches, but stabilized spatial variability across the metacommunity in mesocosms of seagrass. Staddon et al., (2010) quantified the effect of movement corridors on BEF in replicate four-patch metacommunities of moss microarthropods. The absence of corridors led to the extinction of apex predators, increasing prey species abundance. This trophic cascade significantly altered carbon and nitrogen fluxes in isolated habitats. Local extinctions and disruption of ecosystem processes (CO₂ flux, DOC, and total nitrogen in leachate) were mitigated, and even reversed, by the presence of corridors because consumer movement was maintained.

4. SYNTHESIS

Our review of theories and empirical research shows that BEF relationships are dynamic and scale dependent, even in the most human-controlled experiments. Here we discuss four directions for theoretical and empirical synthesis that could guide future work in the near term.

4.1 Varying grain and extent to ‘unveil’ the BEF relationship across scales

As we increase the scale of our analysis we can expect to ‘unveil’ nonlinearity in the magnitude and direction of the BEF relationship (Fig. 2). This can be studied empirically by fixing spatial extent and varying the grain by aggregating units of observation to see how coarse-graining alters the magnitude and stability of BEF over metacommunities with varying levels of connectivity. Alternatively, one can estimate the change in BEF relationship across unconnected systems of varying spatial extent that may differ in species pool size and internal heterogeneity (e.g. across
oceanic islands of different sizes, or ‘islands’ of habitat fragments or lakes). Both address different and important generating mechanisms of scale dependence in BEF and may support distinct scaling relationships because of the role of β-diversity (Mori et al. 2018). For example, we should expect β-diversity and the rate of species turnover to be greater across island systems (Wardle et al. 1997), compared to an equivalent total area from samples of a spatially contiguous system. The greater regional complementarity caused by lower connectivity across island systems should result in stronger scale dependence in BEF (Bond & Chase 2002; Thompson et al. 2018).

So far, theory (Wang et al. 2017, Thompson et al. 2018) and empirical studies (e.g. Barnes et al. 2016) have predominantly focused on revealing changes in stability and BEF as grain is aggregated. Recent analyses (Barry et al. 2019) combine knowledge of species-area and biomass-area relationships to upscale the species richness-biomass relationships. Empirical research has addressed how biodiversity and ecosystem functioning vary across oceanic islands (Wardle et al. 1997) or habitat islands (Gonzalez & Chaneton 2002), but neither addressed BEF as a function of extent. Research on the IAR could be easily extended to cross island comparisons and thereby allow an assessment of the effects of grain and extent. Here, there is an opportunity to link ecosystem functioning to interaction network structure and diversity as it is constrained by habitat space/volume (Post et al. 2000; Tunney et al. 2012).

4.2 Drivers of asynchrony link stability and biodiversity and ecosystem function across scales

Synchrony, within and among species and functional groups, is predicted to affect the magnitude and stability of ecosystem functioning at different scales. The geography of synchrony (Walter et al. 2017) will affect the geography of BEF. Examples include changes in forest growth synchrony...
and carbon dynamics in Eurasian regions due to climate warming and variability (Shestakova et al. 2016), and the global geography and temporal scales of (a)synchrony in primary production on land (Defriez & Reuman 2017) and in the oceans (McGinty et al. 2012; Defriez et al. 2016). Synchrony in EF across scales is influenced by environmental forcing due to changes in the fluxes of energy (irradiance, heat, wind) and matter (nutrient inputs, biomass inputs), and indirectly by the spectrum of frequencies of response and growth within diverse assemblages of species and functional groups (Vogt et al. 2011).

The theory we reviewed shows that synchrony is central to spatial and temporal complementarity and thus the scaling of BEF and stability. For example, the change in IAR with spatial extent is explained by the distance-dependent decay in synchrony in growth dynamics (Wang et al. 2017; Delsol et al. 2018) that is altered by local and regional environmental variability, and turnover in diversity among assemblages (Thibaut and Connolly 2013; Wang et al. 2019). Ultimately, the magnitude and stability of ecosystem functioning at different scales depends upon whether environmental conditions, movement and trophic interactions synchronize or desynchronize species or functional groups (Ives et al. 2000; Gonzalez & Loreau 2009; Gouhier et al. 2010; Wang et al. 2017; Lamy et al. 2019).

Future theory and empirical research could assess three predictors of interspecific synchrony. The first is forcing caused by fluxes in energy which varies considerably across scales on land and in the oceans (Carrara & Vázquez 2010; Vogt et al. 2011; Acevedo-Trejos et al. 2018) and is known to be synchronizing when driven by strong periodic cycles (Blauw et al. 2018), disturbances (Keitt 2008) and autocorrelated random fluctuations (Petchey et al. 1997). This environmental variation
engenders different compensatory responses among species or functional groups at different scales, reflecting variation in seasonal and interannual phenology (Thackeray et al. 2010; Lasky et al. 2016) and asynchronous population fluctuations across trophic levels (Fontaine & Gonzalez 2005; Keitt & Fischer 2006; Vasseur & Gaedke 2007; Loreau & de Mazancourt 2008; Fauchald et al. 2011; Vasseur et al. 2014; Sheppard et al. 2019). Second, movement and connectivity across scales can synchronize population fluctuations, even when separated by great distances. This non-local action can arise from long-distance migration events (Bauer & Hoye 2014) and the topology of environmental teleconnections (Boers et al. 2019). Third, spatio-temporal synchrony can be driven by interactions, such as predation, that occur over a range of scales to couple the dynamics of spatially distinct food webs (McCann et al. 2005; Gouhier et al. 2010). Experiments and empirical surveys that combine two or more of these synchronizing factors should elicit scale dependence in BEF magnitude and stability.

4.3 New theory to tackle the question of scale in BEF

At the most abstract level scale-dependence in BEF relationships can be seen as the outcome of collective dynamics of species persisting as networks of interacting coupled nonlinear oscillators (Kouvaris et al. 2010). Of particular interest is the transition from spatiotemporal disorder to synchronized dynamical regimes, which arise by external forcing, endogenous feedback and feedforward mechanisms, and spatial flows of information. In this context, the search for linear correlations gives way to nonlinear correlations characterized by synchronization and phase coherence among the fluctuations components (Gans et al. 2009; Gouhier et al. 2010). Complementarity among species may occur at one spatial or temporal scale but not at others, depending on the frequency of fluctuations, such as temperature and precipitation, that can drive
population fluctuations and set the productivity of the ecosystem they are embedded within. If causal links between biodiversity and ecosystem function are dynamic and scale-dependent in this way, then theory that formalizes communities as networks of interacting oscillators will be needed. BEF theory can be framed in this way (Chesson et al. 2001; McCann et al. 2005; Gravel et al. 2016; Wang et al. 2019). We now identify several opportunities for further research in this direction.

First, we need theory that treats BEF relationships as dynamic in space and time (Massol et al. 2011a; Leibold et al. 2017). Dynamic BEF relationships result from changes in the diversity (i.e., number, evenness and heterogeneity) of the fluctuating component populations and their interactions (Miele et al. 2019). Complementarity across scales will arise from the scale dependence in the spatial network of interactions (Peterson et al. 1998; Gross & Cardinale 2007; Peters et al. 2007). Reaction-diffusion models of ecosystem patchiness have addressed the emergence of patchiness of species and biomass (van de Koppel et al. 2012; Tarnita et al. 2017), but not the emergence of cross-correlations (or cross-coherence) between biodiversity and ecosystem function at different scales. A cross-scale theory for BEF can be achieved by combining insights from metaecosystem and metacommunity theory (Massol et al. 2011a; Marleau & Guichard 2014; Thompson et al. 2017). Integration of these theories allows a simultaneous analysis of how biodiversity change at different trophic levels affect ecosystem processes within and among patches. The generalized Lotka-Volterra framework suggested by Massol et al. (2011a) although a big step towards integration, assumes direct interactions (e.g. predator-prey interactions) are operating on the same time scale, are not spatially explicit, and do not track the productivity and location of abiotic resources. These assumptions should be relaxed to study scale-
dependence in BEF as emergent from cross-scales interactions among levels of organization in multiplex networks (Scotti et al. 2013; Kéfi et al. 2016; Pilosof et al. 2017).

Moving from landscapes to entire regions or biomes requires models that bridge BEF theory and biogeography (Peters et al. 2008). Functional biogeography links functional diversity, mediated by trait-environment relationships, to major cycles and fluxes in ecosystem function, as constrained by climate gradients (Enquist et al. 2007, 2015; Reichstein et al. 2014). These models have been developed for plants assemblages, but extensions to include other trophic groups such as marine fish assemblages are available (Frainer et al. 2017). Trait-based approaches are already developed for metacommunity models, so there is an opportunity to extend these models to include realistic trait-environment relationships, and to assess how connectivity leads to the correlation between biodiversity and ecosystem function at regional and global scales (Massol et al. 2011b; Reichstein et al. 2014; Garcia et al. 2016).

4.4 Linking theory to new observational data on biodiversity change and ecosystem function

Tests of the theory we have reviewed here will require scale-explicit multivariate data amenable to more sophisticated statistical methods that can assess scale-dependence in BEF relationships. For that, we need multiscale measures of ecosystem processes (Soranno et al. 2019) and biodiversity change (Barnes et al. 2016; Chase et al. 2019). For measuring biodiversity change at different scales, BEF research must harness current methodological developments (Bush et al. 2017), like metagenomics, eDNA (Cristescu & Hebert 2018), and remote sensing (Pau & Dee 2016; Rocchini et al. 2018) and multi-site monitoring networks and experiments. Scale-explicit analyses will require multiscale statistical methods, such as generalized dissimilarity modeling.
(Ferrier et al. 2007), that can be used to predict spatial patterns of turnover in diversity that are crucial to understanding how BEF relationship will change across large spatial and temporal extents (Leibold et al. 2017; Hu et al. 2018; Mori et al. 2018). Integrative data analyses using structural equation modeling can evaluate how BEF relationships might change with scale (Grace et al. 2014, 2016). The structure of these models can include uncontrolled factors that covary with diversity and functioning that are inherent to observational data, especially at large spatial scales. Indeed, structural equation models that incorporates multiple scales have already shown their value here (Barnes et al. 2016; Grace et al. 2016). These approaches can be used in conjunction with frameworks designed for causal inference (Rubin 2005; Pearl 2009) to address multiple causes of change in BEF relationships as we cross scales.

Human impacts on the climate are now so widespread that they can drive patterns of synchrony across large spatial scales (Frank et al. 2016), which as we have discussed governs the scaling of ecosystem functioning and its stability. New multivariate methods (Mahecha et al. 2019) capable of revealing nonstationary interactions among species assemblages, ecosystem processes and climate forcing could be applied to evaluate how BEF effects are changing under climate change. With larger data sets, including time series across a network of spatial locations, methods such as wavelet analysis can be applied to characterize scales of synchrony and cross-coherence between biodiversity change and ecosystem functions. Given good time series these methods can also detect the effects of changing synchrony in the dynamics of species’ fluctuations at different scales (Baldocchi et al. 2001; Keitt & Fischer 2006; Vasseur & Gaedke 2007; Stoy et al. 2009; Cazelles et al. 2014; Walter et al. 2017).
5. HUMAN IMPACTS ON BEF ACROSS SCALES

A major motivation for this synthesis is a need for a deeper understanding of the cross-scale impacts of humans on BEF relationships. This is an imperative next step for BEF research because very little of the Earth’s land surface is now unaffected by humans, with recent estimates indicating that ~52% is now in a state of intermediate modification, and 84% now affected by multiple impacts (Kennedy et al. 2019). Humans modify landscapes by clearing land to make way for agriculture or urban growth, and by altering natural patterns and scales of environmental heterogeneity and disturbance. These effects alter patterns of diversity change locally and regionally (Haddad et al. 2015; Newbold et al. 2015), that provoke extinction debts, invasion and turnover (Kuussaari et al. 2009; Jackson & Sax 2010; Ewers et al. 2013). The effects of human land use on BEF relationships and ecosystem services have already been reported at landscape scales (Mitchell et al. 2014, 2015; Qiu et al. 2018; Winfree et al. 2018; van der Plas et al. 2019).

However, a systematic assessment of how humans affect BEF across scales is needed.

The theory we have reviewed here may explain the impacts of human land use change on the scaling of BEF relationship. We have seen that BEF relationships are sensitive to altered patterns of species turnover in space and time (Keitt & Fischer 2006) because turnover affects the scales at which complementarity to changing environmental conditions is observed. Humans, by fragmenting the landscape, create spatial networks of habitat patches connected to diminishing degrees (Fig. 5A). These alterations in patch connectedness can modify how the BEF slope relates to sampling extent (Fig. 5B), creating complex and unexpected changes in the BEF scaling relationships (Fig. 5C). Empirical verification of such dynamic landscape models will become

Future experiments should address the underlying causes of scale dependence in human dominated landscapes. This can be done by examining how changes in the composition, configuration and connectivity of ecosystem fragments can affect ecosystem functioning and stability (Thompson et al. 2017). This may happen for at least three reasons that may be tested as complementary hypotheses. First, land conversion decreases the total area of available natural habitat (e.g., forests, grasslands). For instance, land ownership patterns can determine the sizes and shapes of remnant habitat patches (Keitt et al. 1997). Across many small patches, $\beta$-diversity can be increased due to random sampling of species occurrences and stochastic extinctions reducing compositional similarity from patch to patch. Alternatively, predictable extinction sequences (e.g. ordered by body size) in small remnant patches can homogenize localities across a large region. Second, habitat loss due to land conversion affects the configuration of remaining patches in a landscape, affecting habitat connectivity (Lamy et al. 2016), which we expect to drive scenarios of extinction mediated by trade-offs in dispersal capacity and sensitivity to environmental conditions in fragmented landscapes. This is predicted to increase turnover and $\beta$-diversity (Germain et al. 2017). Indeed, structural connectivity has been found to alter the slope of the relationship between above-ground carbon and tree functional diversity in remnant forest fragments embedded in crop land (Ziter et al. 2013). Third, humans may impact ecosystem function at larger scales by lowering $\beta$-diversity (biotic homogenization) within regions (Nowakowski et al. 2018), as well as by causing spillover of nonendemic species into adjacent natural habitat (Bell & Tylianakis 2016). Tests of these expectations can be done in experimental landscapes that control patterns of habitat
loss, fragmentation and connectivity (Staddon et al. 2010; Lindo et al. 2012; Haddad et al. 2015), or in systems where patch-to-patch turnover in diversity and composition is controlled directly and can be sampled at multiple spatial scales (Pasari et al. 2013).

6. CONCLUSIONS

BEF relationships are not constrained to small scales. Much of our evidence for BEF relationships comes from small scales (Fig. 1) because that is where we have sought them. However, like many processes in ecology, we expect BEF relationships to span multiple scales, and because of cross-scale feedbacks, their strength and form will change across scales. We reviewed multiple theories (Table 1) that lead us to expect change in diversity to causally drive variation in ecosystem functioning far beyond the local scale. We have suggested in section 4 how progress can be achieved. We need stronger links between scaling theory, spatially explicit models of species networks \textit{that link functional diversity to (a)synchronous patterns of biomass variation that characterize the change in selection and complementarity effects of diversity on ecosystem processes in space and time. A new generation of networked experiments, surveys and remote-sensing observations are needed to inform global ecosystem models that incorporate BEF knowledge (Enquist et al. 2003; Ward et al. 2012; Harfoot et al. 2014; Asner & Martin 2016; Lausch et al. 2016; Acevedo-Trejos et al. 2018; Schweiger et al. 2018; Tréguer et al. 2018). These connections must be understood if BEF research is to foster progress towards the UN’s Sustainable Development Goals and our efforts to manage biodiversity for the many benefits ecosystems provide people from local to global scales (Dee et al. 2017; Isbell et al. 2018).

Acknowledgements
We thank Mary O’Connor and Jane Cowles for valuable discussions and suggestions for improving the manuscript. AG acknowledges the support of NSERC, Killam Fellowship, the Quebec Centre for Biodiversity Science (QCBS) and the Liber Ero Chair in Biodiversity Conservation. DS acknowledges support of NSERC. RMG was supported by the Biodiversity Research Centre and Killam trust, and NSERC. JC, LD, and FI acknowledge support from the US National Science Foundation’s LTER Network Communications Office (DEB-1545288). PLT was supported by NSERC and the Killam trust. SK was supported by the ANR project ARSENIC (ANR-14-CE02-0012). YZ, ML and JMM were supported by the TULIP Laboratory of Excellence (ANR-10-LABX-41), and YZ and ML by the BIOSTASES Advanced Grant (grant agreement no. 666971), and JMM by the FRAGCLIM Consolidator Grant (grant agreement no. 726176) from the European Research Council under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 666971). This paper arose from a joint working group supported by QCBS (working group #12) and the Centre for Biodiversity Theory and Modelling (Moulis, France).

Bibliography


Table 1. Expected effects of (a) spatial, (b) temporal and (c) organizational scale on BEF relationships derived from the theoretical and empirical studies we reviewed. The effects discussed in the main text are in boldface.

<table>
<thead>
<tr>
<th>Process change as we increase spatial scale</th>
<th>Magnitude of function</th>
<th>Stability of function</th>
<th>Spatial scale at which process applies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical properties of aggregation</td>
<td>BEF slope increases as spatial grain becomes coarser due to non-linear averaging of different finer-grain BEF relationships, as well as spatial variation in fine-grain diversity (Thompson et al. 2018), coupled with evidence for spatial variation in fine-grain BEF slope (Sullivan et al. 2017, Vilà et al. 2013, Liang et al. 2016)</td>
<td>Observation error associated with sampling biodiversity distributions can be a component of biodiversity-stability relationships and will increase with spatial scale (due to increased environmental heterogeneity and patterns of rarity) and decrease with sample size due to averaging over multiple observations (Mazancourt et al. 2013).</td>
<td>Patch to continental</td>
</tr>
<tr>
<td>Spatial turnover in species composition increases, due to drift, dispersal limitation or species sorting</td>
<td>As spatial grain increases, BEF relationship steepens when new species are encountered across space, and erodes when all species have already been encountered (Thompson et al. 2018) Landscapes require more species to maintain functioning than sites because functionally important species differ between sites (Winfree et al. 2018; Lefcheck et al. 2019)</td>
<td>Habitat selection by mobile organisms can alter effects of diversity on stability of functions (France and Duffy 2003)</td>
<td>Patch to landscapes to regions</td>
</tr>
<tr>
<td>Heterogeneity and range of environmental</td>
<td>BEF strengthens because of increased expression of niche complementarity</td>
<td>Insurance hypothesis, both local and spatial, can be explained via temporal and spatial niche</td>
<td>Micro-habitat to patch</td>
</tr>
<tr>
<td>conditions increases</td>
<td>(Dimitrakopoulos &amp; Schmid 2004; Tylianakis et al. 2008). Direct effects of environment on ecosystem functions reduces relative importance of biodiversity (e.g. Srivastava &amp; Vellend 2005)</td>
<td>complementarity. Both predict stabilizing effects of increasing diversity, and these benefits become more apparent with increasing spatial and temporal scales.</td>
<td>Patch to landscape to region</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td>Dispersal influences local dynamics</td>
<td>In combination with a spatially and temporally heterogeneous environment, moderate dispersal can permit species to efficiently track spatial change in their optimal environment, increasing function (Loreau et al. 2003, Thompson et al. 2017)</td>
<td>Rates of movement from patch to patch by mobile consumers can either stabilize (Loreau et al. 2003) or destabilize ecosystem function (Marleau et al. 2014).</td>
<td>Patch to region</td>
</tr>
<tr>
<td>Potential for spatial asynchrony in local population dynamics</td>
<td>Increasing asynchrony predicts an increase in average EF.</td>
<td>In combination with a spatially and temporally heterogeneous environment, asynchrony in population dynamics at the local level results in stabilization at larger spatial scales (Loreau et al. 2003, Wang and Loreau 2016)</td>
<td>Patch to region</td>
</tr>
<tr>
<td>Spatial coupling of functions between habitats</td>
<td>As spatial scale increases, ecosystem functions include energy and matter flow between habitats. Thus, BEF effects in one habitat may support ecosystem functions in a connected habitat (Alsterberg et al. 2017).</td>
<td></td>
<td>Single habitat/ecosystems to multiple habitats/ecosystems</td>
</tr>
<tr>
<td>Allow feedbacks from EF -&gt; B</td>
<td>Ecosystem functions enabled by one group of species provide opportunities for other species through niche construction, processing chains and autocatalytic cycles, food webs, facilitation networks and cross-</td>
<td>Temporal variance in ecosystem functions affects the persistence of species; intermediate levels of variability often promote diversity (e.g. creating temporal niches, disrupting competitive</td>
<td>Patch to landscape</td>
</tr>
</tbody>
</table>


**Increasing temporal scale**

<table>
<thead>
<tr>
<th>Magnitude of function</th>
<th>Stability of function</th>
<th>Temporal scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>BEF slope changes with successional stage (Reich <em>et al.</em> 2012; Lasky <em>et al.</em> 2014; Guerrero-Ramírez <em>et al.</em> 2017), and aggregation of variable BEF slope is subject to non-linear averaging (a temporal analog to Thompson <em>et al.</em> 2018)</td>
<td>Ability to detect temporal complementarity in population dynamics, a main mechanism underlying diversity-functional stability relationships, increases with observation time (Loreau &amp; Mazancourt 2013)</td>
<td>One to many generations</td>
</tr>
<tr>
<td>Diverse communities are able to accumulate limiting nutrients within ecosystem components, fueling further growth (Reich <em>et al.</em> 2012)</td>
<td>Ability to detect temporal complementarity in population dynamics, a main mechanism underlying diversity-functional stability relationships, increases with observation time (Loreau &amp; Mazancourt 2013)</td>
<td>One to many generations</td>
</tr>
<tr>
<td>In assembling communities, BEF relationships strengthen due to increasing strength of complementarity (Cardinale <em>et al.</em> 2007). Other species compensate for lost species through vegetative growth or colonization, reducing influence of biodiversity on function (Kardol <em>et al.</em> 2018)</td>
<td>Ability to detect temporal complementarity in population dynamics, a main mechanism underlying diversity-functional stability relationships, increases with observation time (Loreau &amp; Mazancourt 2013)</td>
<td>One to many generations</td>
</tr>
<tr>
<td>Dee et al. (2016) found evidence for a performance-effect of functional diversity, by buffering fisheries yields against within-year temperature variability.</td>
<td>Diversity can stabilize functions in the face of the destabilizing effects of reddened environmental noise (Gonzalez and De Feo 2007)</td>
<td>Years to decades</td>
</tr>
</tbody>
</table>

**Increase organizational scale**

<table>
<thead>
<tr>
<th>Magnitude of function</th>
<th>Stability of function</th>
<th>Organizational scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased rates of ecosystem functions, due to niche complementarity between genotypes</td>
<td>Increased functional resistance to disturbance (Hughes &amp; Stachowicz 2004)</td>
<td>From individuals to population</td>
</tr>
<tr>
<td>Include multiple populations of a species</td>
<td>Variability in function is reduced when populations have independent or negatively covarying dynamics (portfolio effect, e.g. (Schindler et al. 2010))</td>
<td>Single population to metapopulation</td>
</tr>
<tr>
<td>Include multiple trophic levels of a food web.</td>
<td>Higher trophic levels can change BEF effects at lower levels, e.g. by altering the relative abundances and interaction strengths of lower-level species or directly influencing function (Worm &amp; Duffy 2003)</td>
<td>Higher trophic levels can alter the relationship between diversity and ecosystem stability, depending on the strength of the trophic interactions (Thébault &amp; Loreau 2005; Jiang et al. 2009)</td>
</tr>
<tr>
<td>Include multiple communities of a metacommunity</td>
<td>See all “patch to landscape” or “patch to region” entries in spatial scale table</td>
<td>See all “patch to landscape” or “patch to region” entries in spatial scale table</td>
</tr>
<tr>
<td>Include multiple habitat or ecosystem types of a meta-ecosystem</td>
<td>Habitat diversity can have strong impacts on ecosystem functioning when habitats complement each other in the types of energy and elemental processing (Alsterberg et al. 2017)</td>
<td></td>
</tr>
</tbody>
</table>
Figure captions

Figure 1. The three dimensions of scale in BEF research: time, space and organization (see Box 1 for definitions). Most empirical studies in BEF (shown by black dots) fall within a constrained volume of this scale box: days to weeks in the case of micro and mesocosm experiments, and years to two decades in the case of some grassland and forest diversity experiments. The size of most experimental plots is typically less than a hectare, although the spatial extent of the largest experiment was continental (BIODEPTH). Empirical studies could sample larger scales of variation by combining data from remote sensing technologies, in situ probes and buoys, surveys using long transects, and geographic networks of replicated experiments with controlled perturbations at different scales, deployed for multiple years and over broad spatial extents to capture shifting gradients of environmental heterogeneity.

Figure 2. The scale of observation directly affects biodiversity and ecosystem functioning but also affects the relationship between them. In panel A assemblages are sampled across a spatial or temporal gradient in the environment (Env), species occurrence by trophic level (with corresponding food web shown, right), diversity measured as richness (sum of species occurring at each location) and ecosystem functioning (e.g. productivity, or total community flux). These samples (scale 1) can be aggregated over space or time (scale 2). Diversity and ecosystem function each show characteristic scaling relationships with increasing spatial or temporal extent (B and C respectively), and the difference in these scaling relationships contributes to a scale-dependence in the BEF relationship (D) which can be projected as a BEF slope by scale plot (E). With just two scales in this example, we have only incomplete sampling of the potentially nonlinear BEF slope by scale relationship (blue line in E).
Figure 3. Scale dependence in BEF relationships. (A) Expectations from theory on scaling of biodiversity and ecosystem functioning via aggregation across sites (redrawn after Thompson et al. 2018). The strength of biodiversity effects, $b_A$, as measured by the slope of the BEF relationship at different spatial scales when there is incomplete compositional turnover across local patches (see main text). Different degrees of compositional turnover are indicated by the different coloured lines (low values of $B_1$ correspond to low turnover; $B_1 = 0$ indicates complete turnover).

(B) Expectations from theory on scaling of biodiversity and ecosystem functioning with area (redrawn after Delsol et al. 2018). The green line shows the expected SAR (species-area relationship), and a similar blue curve for the IAR (invariance-area relationship). The grey area shows the set of possible SARs and IARs for a fixed configuration of individuals. Its upper boundary (red) coincides with the expected proportional relationship between the number of individuals and area. For SARs, this boundary corresponds to the limiting case where each individual belongs to a different species, yielding a linear SAR. For IARs, it corresponds to the limiting case where all individuals have independent productivity fluctuations, yielding a linear IAR. Productivity is expected to scale proportionally to the number of individuals, and thus to follow the same linear relationship. Bottom panels a, b and c show the distributions of individuals (a and b) and species ranges (b and c) at three spatial scales. Individuals and ranges of different species are indicated by different colours.

Figure 4. Temporal environmental autocorrelation alters the scale-dependence in species fluctuations, mean EF (orange line) and stability (redrawn after Gonzalez and DeFeo 2007). The top panels (A-C) show increasing autocorrelation in the environmental fluctuations (shown in grey) from left to right. RAR = Resource Assimilation Rate: (A) white noise, with no
autocorrelation, (B) $1/f$ noise, or pink noise, where the power spectral density of the environmental fluctuations is inversely proportional to the frequencies $f$ composing the signal. (C) $1/f^2$ environmental fluctuations (red noise). The population dynamics for 3-species resource competition. The species have distinct, but overlapping, environmental niches (left-hand side of each of the top panels) which are shown as coloured Gaussian curves. The mean ecosystem function (orange) and species’ fluctuations are dominated by low frequency fluctuations as the environment becomes more autocorrelated (A to C). Panels D and E, respectively, show how mean community biomass and community stability change as a function of species richness (2-24 species) and the degree of autocorrelation characterized by the slope of the exponent (eight levels 0-2).

**Figure 5.** Right: Satellite image of an agricultural landscape with remnant forest fragments. Left: Predictions for the change in BEF slope as the scale of observation increases for three landscapes with varying degrees of fragmentation (simulated data). Top row: Stylized landscape patterns with different fragmentation patterns of forest habitat (dark green) cleared agriculture (white background): (a) homogeneous forest (${x =}$ northing, ${y =}$ easting), (b) fragments with varying diversity and productivity, with links indicating connectivity by seed dispersal, (c) isolated fragments with lower average diversity and productivity and fewer links. At each scale of observation, denoted by the coloured sampling windows in (a-c), species richness and productivity are measured at different locations across a landscape by sliding the window. Middle row: Change in the linear relationship between species richness and productivity at different scales of observation for each landscape type (d-f). Each coloured line is composed of measurements of species richness and productivity from multiple windows at a given scale. Species richness and
productivity increases with the spatial scale of observation for all three landscape types but the form of the BEF relationship varies. Bottom row: Change in the BEF slope as a function of the scale of observation for each landscape type (h-j). Each point corresponds to the value of the slope of the line of same colour in the respective above figure. At a small sampling scale (orange window in (a)) the BEF slope is low and similar in all three landscape types (orange points in (h-j)). At that scale, species richness and productivity are small and not affected by fragmentation (orange lines in (d-f)). At an intermediate sampling scale (red window in (a)), the BEF slope increases in all three landscape types. At that scale, sampling windows accounted for more species richness and higher level of productivity leading to stronger BEF effects. While fragmentation has reduced both biodiversity and productivity (red lines in (d-f)), no notable impact on the BEF slope is observed at this scale (red points (h-j)). At a large sampling scale (blue) the BEF slope decreases in the homogeneous landscape (a, d, h) since most species have already been sampled producing no additional biodiversity effects on productivity. However, when fragments are isolated (c), even if species richness and productivity are lower (f), a wide range of species richness and productivity are sampled (blue line in (f)) leading to an increase in BEF slope (blue point in (j)). The effect of species turnover on the BEF slope is also observed, although to a lesser degree, in the landscape with linked fragments (b, e, i) since species turnover is reduced by the ability of species to disperse across the landscape. At a very large sampling scale (green window in a) the BEF slopes decreases in all three landscape types but at different levels (green points in (h-j)). While productivity is higher at that scale, species richness is similar in all sampling windows (green lines in (h-j)).
Figure 2

A

Env.

Species occurrence

Diversity

Ecosystem functioning

Scale 1: Sample scale

Scale 2: Aggregation

B

Log(diversity)

Log (extent)

C

Log (EF)

Log (extent)

D

Ecosystem functioning

E

BEF slope

Unobserved scales

scale\textsubscript{aggregation}
Figure 3

A

B

(a)

(b)

(c)

100 m x 100 m

6 km x 6 km

400 km x 400 km

$B_1$

-0.75

-0.5

-0.25

-0.1

-0.05

-0.025

-0.01

0
Figure 4

A

B

C

D

E

community biomass

community stability

Environment color
color

number of species

number of species
Figure 5

![Diagram showing various plots and graphs related to productivity, species richness, and BEF slope across different spatial scales.](image)
Box 1: Features of scale necessary for multiscale BEF research

Time, space, and ecological organization are three important dimensions over which the BEF relationship vary (Fig. 1). Each dimension has three fundamental aspects of scale: 1) intrinsic process scales over which B and EF vary and covary, and 2) observation scale(s), defining how the system is partitioned (e.g. size of experimental unit) and sampled in space and time (Fig. 2), and 3) analytical scale defining the scales over which inferences are made (Dungan et al. 2002). The intrinsic scales are expressed in biologically relevant dimensions of space and time (i.e. generations or frequencies of (a)periodic ecosystem function in time or space); observation scales are characterized by their ‘grain’, the time or space resolution of individual samples, and ‘extent’, the scale encompassing all observations. Analytic scales relate to how data are aggregated and transformed to optimize model fitting and inference.

Sampling governs which intrinsic scales are observed and how well they are sampled (Fig. 2). Under sampling can either mask or bias our estimate of the relationships between biodiversity and ecosystem function. In the absence of a strong a priori expectation for the scale(s) at which BEF interactions are strongest, multiscale sampling is required to capture ranges of variation in biodiversity and ecosystem function known to occur in the system; 3) the scale(s) of analysis and inference chosen to model BEF. This may involve a microscopic approach focusing on variation among individuals and their metabolic activities, a mesoscopic approach that examines how patchiness biodiversity and ecosystem functioning determine the BEF relationship, or alternatively, a macroscopic approach whereby measures of diversity (e.g. entropy, functional trait distributions) are used to summarize variation across many assemblages and used to predict ecosystem function, such as NPP or carbon fluxes at biogeographic scales.
Temporal scale: temporal dimensions of observation, where the duration of a single measurement sets its resolution (seconds, days), the frequency defines the time elapsed between measurements (e.g., annual), and the interval between the first and last measurement sets the temporal extent. In general, a signal can only be reconstructed from its samples if it is sampled at least twice as fast as its highest frequency component. The intrinsic time scales of biodiversity include the rates of temporal species turnover arising from colonization and extinction, and the fluctuations (e.g., variance, extremes) in ecosystem function (energy flux, or biomass production), that may have a strong seasonal (e.g. annual) and meso- and macroclimatic periodicities (e.g., ENSO or El Niño), and a multiscale random component (i.e., environmental noise) with autocorrelation.

Spatial scale: the spatial dimensions of the study, the extent and grain of the study area (e.g. plot size) and sampling effort (e.g. spatial coverage). Again, we also include the intrinsic scales of diversity turnover in space and distance decay in similarity. The latter leads naturally to the notion of effective community diversity, and β-diversity that links local (α) diversity at the grain studied, to regional (γ) diversity at the maximum extent studied.

Organizational scale: measures of non-randomness in biodiversity (taxonomic, functional or phylogenetic diversity and their effective numbers), and relational measures of organization characterizing species’ associations and interactions (e.g. food web, or mutualist networks) that scale in space and time. These include trophic complementarity and the vertical/horizontal diversity of food webs. We also mean spatial network organization, where dispersal and fluxes of resources can link patches to create dynamically coupled assemblages of species (e.g. metacommunities). These organizational scales align when the complexity of interaction networks
has strong spatial and temporal dimensions.

Our review of theory shows that BEF research must vary the range of observation scales if it is to capture the range of intrinsic process scales and make strong statistical inferences about this scale dependence. The multiscale nature of BEF relationships suggest that where possible we must contrast or manipulate aspects of intrinsic scale (e.g., climate variation, nutrient pulses) to reveal strength of scale-dependence in experimental and natural ecosystems.