



Theses and Dissertations

2020-08-03

Analyzing Metacommunity Models with Statistical Variance Partitioning: A Review and Meta-Analysis

Kevin Vieira Lamb
Brigham Young University

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Life Sciences Commons](#)

BYU ScholarsArchive Citation

Lamb, Kevin Vieira, "Analyzing Metacommunity Models with Statistical Variance Partitioning: A Review and Meta-Analysis" (2020). *Theses and Dissertations*. 9248.
<https://scholarsarchive.byu.edu/etd/9248>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

Analyzing Metacommunity Models with Statistical Variance
Partitioning: A Review and Meta-Analysis

Kevin Vieira Lamb

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Master of Science

Russell Rader, Chair
Mark Belk
Jerald Johnson

Department of Biology
Brigham Young University

Copyright © 2020 Kevin Vieira Lamb

All Rights Reserved

ABSTRACT

Analyzing Metacommunity Models with Statistical Variance Partitioning: A Review and Meta-Analysis

Kevin Vieira Lamb
Department of Biology, BYU
Master of Science

The relative importance of deterministic processes versus chance is one of the most important questions in science. We analyze the success of variance partitioning methods used to explain variation in β -diversity and partition it into environmental, spatial, and spatially structured environmental components. We test the hypotheses that 1) the number of environmental descriptors in a study would be positively correlated with the percentage of β -diversity explained by the environment, and that the environment would explain more variation in β -diversity than spatial or shared factors in VP analyses, 2) increasing the complexity of environmental descriptors would help account for more of the total variation in β -diversity, and 3) studies based on functional groups would account for more of the total variation in β -diversity than studies based on taxonomic data. Results show that the amount of unexplained β -diversity is on average 65.6%. There was no evidence showing that the number of environmental descriptors, increased complexity of environmental descriptors, or utilizing functional diversity allowed researchers to account for more variation in β -diversity. We review the characteristics of studies that account for a large percentage of variation in β -diversity as well as explanations for studies that accounted for little variation in β -diversity.

Keywords: community ecology, variance partitioning, environmental, spatial, stochastic, deterministic, meta-analysis

ACKNOWLEDGMENTS

I would like to acknowledge Dr. Russell Rader for his contributions to this research and for his guidance through each stage of the research and writing processes. I would like to thank Nicholas Suiter, Stephanie Suiter, Thomas Knapp, and Jessica Pukahi for their assistance in acquiring and summarizing many of the articles utilized for this research. Lastly, I would like to thank Georgina Lamb for her continual encouragement and support from the start of this project to its completion.

TABLE OF CONTENTS

Title Page	i
ABSTRACT.....	ii
ACKNOWLEDGMENTS	iii
TABLE OF CONTENTS.....	iv
LIST OF TABLES	v
LIST OF FIGURES	vi
Introduction.....	1
Methods.....	5
Results.....	8
Discussion.....	11
References.....	19
Tables and Figures	26

LIST OF TABLES

Table 1. Environmental descriptors used in <i>complex</i> studies	26
Table 2. The number of articles and variance partitioning data sets	27
Table 3. Analyses showing percentages for environmental (E), spatial (S), and shared (E+S) components of β -diversity that were outliers.....	28
Table 4. Tukey-Kramer post-hoc pairwise comparisons between the mean percentage of the total β -diversity attributed to environmental, spatial, and shared components in variance partitioning analyses	30
Table 5. ANOVA between <i>simple</i> versus <i>complex</i> data sets	31
Table 6. ANOVA between studies using taxonomic identifications versus functional groups	32

LIST OF FIGURES

Figure 1. Number of peer-reviewed articles that partitioned the variation in β -diversity into 4 additive components per year and by ecosystem type	33
Figure 2. Box plots of data sets collected prior to 2020 showing the variation around the average unexplained percentage of β -diversity, and the average percentage of β -diversity attributed to the environment, spatial factors, and the shared components.....	34
Figure 3. Scatter plots showing the relationship between the number of environmental descriptors and the percentage of unexplained variation and environmental variation.....	35
Figure 4. Box plots showing the percentage of the total variation in β -diversity attributed to unexplained, environmental, and shared components of variance partitioning analyses between studies using <i>simple</i> versus <i>complex</i> environmental descriptors	36
Figure 5. Box plots showing the percentage of the total variation in β -diversity attributed to unexplained, environmental, spatial, and shared components of variance partitioning analyses comparing studies using taxonomic diversity versus functional diversity	37

Introduction

The relative importance of deterministic processes versus chance is one of the most important questions in science because it affects our ability to predict future states (Rosenberg & McShea 2008; Vasudevan 2018). Unpredictable variation in nature (uncertainty) is the outcome of randomness and an imperfect knowledge of deterministic processes (Pielou 1972; Landsman 2016). When nature is governed by deterministic processes, similar conditions produce consistent patterns across space and time, and we can predict the future state of a system given sufficient knowledge of its history and current properties. If stochastic forces dominate nature, then randomness increases, and our ability to predict future states decreases.

The metacommunity perspective is the most recent attempt to determine the diversity of local communities (Wilson 1992). Metacommunity theory explains how dispersal interacts with selective forces at various scales (local to global) to determine the distribution of species (Leibold *et al.* 2004; Vellend 2016). Factors that determine the distribution of species and how they assemble to form local communities will determine the dis-similarity in the species composition across localities in a region, which is measured by β -diversity (Whittaker 1960).

Metacommunity theory originally proposed 4 models on how species assemble to form a local community (Holyoak *et al.* 2005). *Environmental filtering* assumes that all species could disperse to all localities across a region but they can only colonize local sites that meet their niche requirements. Extinctions are deterministic as species cannot persist at sites that do not meet their niche requirements (Smith 1989; Jablonski 2017). *Neutrality* assumed that species have equal fitness and thus, respond similarly to the environmental differences among sites. By chance, they are absent at sites that they could otherwise inhabit because of dispersal limitations (Holyoak & Loreau 2006; Guichard 2017). Extinctions are random because no species can

mount a sufficient fitness advantage to drive other species extinct (Hubbell 2001). *Source-sink* dynamics (*mass effects*) represent species that persist where they are poorly adapted (sinks) because of the frequent dispersal of individuals from source sites in which they are well-adapted (Kadmon & Tielbörger 1999). According to the *patch dynamics* model, the presence of species in a local community was determined by a competition/colonization tradeoff. That is, inferior competitors that disperse well (“fugitive species”) were driven locally extinct by superior competitors that poorly disperse (e.g. Hutchinson 1951). Fugitive species are continually on the move driven extinct by superior competitors in one locality as they colonize another. (Winegardner *et al.* 2012) proposed that *source-sink* dynamics and *patch dynamics* are special cases of environmental filtering because mass effects assumes that species are deterministically sorted into high quality sites where they are best fit (sources) and low quality sites where they are poorly fit (sinks). Similarly, *patch-dynamics* assumes differences in the fitness of species driven by strong hierarchical competition, also a deterministic force.

Statistical variance partitioning (VP) is the most recent and most prolific attempt to infer the relative importance of metacommunity dynamics using descriptive data on the species composition of local sites within a metacommunity. Variance partitioning methods separate β -diversity into 4 additive components: 1) pure environmental (E), 2) pure spatial (S), 3) shared spatial and environmental (E+S), and 4) the percentage of unexplained variation in species composition unaccounted for by components 1-3. *Environmental filtering* is deemed important if E accounts for a significant percentage of the total variation in the species composition among sites. As such, β -diversity is correlated with environmental descriptors (Legendre & Legendre 1998; Cottenie 2005). *Neutral* dynamics associated with dispersal limitations are deemed important if the spatial arrangement of sites (S) is correlated with β -diversity. The shared

component (E+S) shows the decay in the similarity of the species composition among sites as distance increases between sites (Bauman *et al.* 2019). A significant E+S can detect the importance of *mass effects* (e.g. Soininen 2007) and/or the importance of spatially structured environmental factors (Bauman *et al.* 2019).

Variance partitioning studies were reviewed by Cottenie (2005) early in the application of this method. In each study, he limited the number of environmental descriptors to 4 in order to facilitate comparisons among studies with a different number of environmental factors (Cottenie 2005). He concluded that the environment accounted for the majority of variation in β -diversity, and he speculated that including more environmental variables would have likely increase the percentage of β -diversity attributed to the environment (Cottenie 2005). The perception that environment filtering accounts for more variation in β -diversity than spatial or shared components has often been repeated in the literature (e.g. Van der Gucht *et al.* 2007; Landeiro *et al.* 2012; Algarte *et al.* 2014).

Soininen gathered VP studies through 2012, including Cottenie's data, to explore the ecological factors (e.g. body size and dispersal capabilities) correlated with both environmental descriptors (Soininen 2014) and spatial factors (Soininen 2016). We found that environmental descriptors in 98% of the VP studies in Cottenie's (2005), and Soininen's (2014 and 2016) reviews consisted of abiotic parameters measured at the local scale (e.g. temperature, water chemistry, soil type). Since then a variety of factors measured at various scales have increased the complexity of environmental descriptors used in VP studies. For example, more recent studies have included species interactions, the presences of exotic species, phylogenetic relatedness of species, the glacial history of a region, topography (e.g. valley form), and human effects (e.g. pollution and land-use practices).

VP analyses based on functional groups that define a species role in the ecosystem should be better correlated with environmental factors and account for more variation in β -diversity than VP analyses based on species (e.g. Gianuca *et al.* 2018). Recent studies have partitioned β -diversity based on species traits (e.g. body size, dispersal ability, feeding strategy) to create a matrix showing functional groups distributed across sites (e.g. Carvalho & Tejerina-Garro 2015). Species will vary across sites depending on colonization, extinction, and speciation dynamics whereas, functional groups can persist across sites as long as a single species within the group remains. Consequently, the percentage of unexplained variation in VP analyses using functional groups should be significantly smaller than VP analyses based on species.

We reviewed the literature on variance partitioning in metacommunities prior to 2020, which added 116 articles that were not included in previous reviews (Cottenie 2005; Soininen 2014; Soininen 2016). We have updated the average variation attributed to environmental, spatial, and shared effects with a focus on how much of the total variation remains unexplained. The amount of unexplained variation is important because it is potentially misleading to draw conclusions about the relative importance of metacommunity dynamics (e.g. *species sorting* versus *neutrality*) if the majority of the variation remains unexplained.

We hypothesized that the number of environmental descriptors in a study would be positively correlated with the percentage of β -diversity explained by the environment, and that the environment would explain more variation in β -diversity than spatial or shared factors in VP analyses (Hypothesis 1). We hypothesized that the percentage of unexplained variation in β -diversity would be lower in this review relative to previous reviews (Cottenie 2005; Soininen 2014; Soininen 2016) because an increase in the complexity of environmental descriptors in recent studies would account for more of the total variation in β -diversity (Hypothesis 2). We

also tested the hypothesis that studies based on functional groups would account for more of the total variation in β -diversity than studies based on taxonomic data (Hypothesis 3).

Methods

Literature Search

We collected 402 articles that used variance partitioning of β -diversity by: 1) searching the “All Databases” section of the Web of Science (see Appendix A for key phrases), 2) including all previous articles reviewed by Cottenie (2005) and Soininen (2014 and 2016), and 3) searching the bibliography of all articles collected in (1) and (2) using EndNote (version X9 build 12062). Of the 402 articles, we eliminated those that did not compare all 4 components of VP analyses (environmental, spatial, shared and unexplained) to explore metacommunity explanations of diversity patterns (e.g. *environmental filtering* versus *dispersal limitations*). For example, studies were eliminated if they used VP to compare the effects of environmental attributes (e.g. fire frequency versus urbanization) instead of comparing metacommunity models (Abdelaal 2017). This left 170 articles and 753 VP data sets. We used the 753 data sets to: 1) test our hypotheses, 2) show the number of VP studies comparing metacommunity models by year and by ecosystem type, and 3) calculate the average percentage of the total β -diversity attributed to environmental, spatial, shared, and unexplained components. Box plots allowed us to show extreme percentages of E, S, E+S and the unexplained components of β -diversity in VP analyses. Outliers were $1.5\times$ greater than the 3rd quantile for E, S, and E+S and $1.5\times$ smaller than the 1st quantile for the percentage of unexplained variation in β -diversity.

Statistical Analyses

To test Hypothesis 1, we calculated simple linear regressions (R Development Core Team 2019) between the number of environmental descriptors used in each data set versus the percentage of unexplained variation in β -diversity, and versus the percentage of β -diversity attributed to the environment (E). We excluded data sets from Cottenie (2005) because he did not report the actual number of environmental predictors in each study. This left 595 data sets used in this analysis. Neither of the regressions required a transformation of the data because both met parametric assumptions of normality and equal variances. Also, we did not calculate the correlation between the number of environmental descriptors and the percentage of variation attributed to spatial factors because the variation attributed to spatial factors is not effected by the number of environmental descriptors.

We used a 1-way ANOVA to determine if there were significant differences in the average percentage of the total variation in β -diversity attributed to environmental (E), spatial (S), and shared (S+E) components of VP analyses based on 753 data sets. We used Tukey-Kramer post-hoc pairwise comparisons among each level of the analysis to test the hypothesis that the environmental component would account for a greater percentage of the total variation in β -diversity than either the spatial or shared components. These data met parametric assumptions of normality and equal variances and thus, did not require a transformation.

We used a MANOVA (R Development Core Team 2019) to determine if there were differences between *simple* versus *complex* VP analyses (Hypothesis 2) in one or more of the correlated dependent variables. The dependent variables were the percentage of the total β -diversity in each data set attributed to the environment (E), the shared component (E+S), and the percentage of unexplained variation. We did not include the spatial component because an increase in the complexity of environmental descriptors should not effect in the percentage of

variation in β -diversity attributed to spatial factors. *Simple* analyses used only local abiotic factors as environmental descriptors, whereas *complex* analyses used local abiotic factors, plus, at least one additional complex environmental descriptor as shown in Table 1. These multivariate data could not be transformed to meet parametric assumptions of homoscedasticity (Box's M test $p < 0.001$) because of unequal sample sizes between *simple* and *complex* analyses. Thus, we used Pillai's trace as the test statistic in the MANOVA because it is more robust than an F-statistic in cases with unequal sample sizes (Ateş *et al.* 2019). We also used separate ANOVA tests (R Development Core Team 2019) on each dependent variable to determine which dependent variables might show differences between *simple* and *complex* analyses. These univariate data were corrected for heteroscedasticity using a natural log transformation (Levene's test $p > 0.01$). We also used a Bonferroni correction (Miller 1981; Allen & Bennett 2008) in the ANOVA tests to reduce the level of significance to $p = 0.0125$ because the *complex* data were based on a small sample size (Table 1).

We again used a MANOVA followed by separate ANOVA tests on the environmental, spatial, shared and unexplained components of β -diversity to determine if analyses based on functional groups differed from analyses using taxonomic identifications (Hypothesis 3). Pillai's trace was used as the MANOVA test statistic (Ateş *et al.* 2019), and the level of significance was set at $p = 0.0125$ after a Bonferroni correction (Miller 1981; Allen & Bennett 2008). A natural log transformation did not correct for heteroscedasticity in the MANOVA (Box's M test $p < 0.001$), but it did correct for heteroscedasticity when an ANOVA was applied separately to each response variable (Levene's test $p > 0.01$). Data sets were excluded from the functional category if VP was performed on species within functional groups rather than using the groups themselves.

Results

Descriptive Statistics

One hundred and sixteen VP articles on metacommunity dynamics have been published since 2012 when Soininen (2014 and 2016) concluded his data collection (Table 2). Variance partitioning procedures (pRDA, db-RDA, CCA, Mantel tests, and regression) continue to be a very common method for examining the relative importance of metacommunity dynamics (Figure 1a). Partial Redundancy Analysis (pRDA) was used in 78% of the VP studies, Canonical Correlation (CCA) in 15%, Mantel tests in 4%, and regression in 3%. Most of the studies were collected in freshwater (rivers, lakes and wetlands) and terrestrial ecosystems (Figure 1b), with only 5 VP articles from the marine environment (Figure 1b). Also, the average percentage of unexplained variation (65.6%) has increased since Cottenie's (2005) review (Table 2). Taken together, the sum of the average percentages of E, S, and E+S only amounted to 34.4% of the total variation in β -diversity (Table 2). There was also considerable variation around each component (Figure 2). That is, unexplained percentages ranged from 2.1% to 100% with 616 analyses out of 753 (82%) reporting $\geq 50\%$ unexplained variation. By contrast, E, S, and E+S accounted for $\geq 50\%$ of β -diversity in only 7, 3, and 3 out of 753 analyses, respectively.

The outliers of VP analyses revealed some surprising results (Table 3). Although the sum of the average percentages of E, S, and E+S only amounted to 34.4% of the total variation in β -diversity, there were 8 out of 753 VP analyses that accounted for $> 84\%$ of the total variation in β -diversity (Table 3). This shows that VP is capable of accounting for the majority of variation in β -diversity. Also, 100% of the studies with extreme environmental or spatial percentages of β -diversity discussed likely explanations, whereas, 41% of the studies with extreme shared percentages did not attempt an explanation. This suggests the difficulty of

interpreting the meaning of a large shared component. Overall, there were no obvious trends in the taxa of the outlier table as the list included microbes of various kinds, multiple types of invertebrates and vertebrates, and a diverse group of plants (Table 3).

Hypotheses 1: the number of environmental descriptors

Contrary to Hypothesis 1, there was no relationship (slope = -0.09; $p = 0.35$; adj $R^2 < 0.01$) between the number of environmental descriptors and the percentage of unexplained variation in β -diversity (Figure 3a). However, there was a significant, albeit weak, positive relationship (slope = 0.19; $p = 0.001$; adj $R^2 = 0.02$) between the number of environmental descriptors and the percentage of β -diversity explained by the environment (Figure 3b).

The perception that environmental filtering is the primary force determining metacommunity dynamics was supported by this review. ANOVA showed significant differences among the 3 components of VP analyses ($F_{2, 2256} = 82.4$; $p < 0.001$). That is, the percentage of the total variation in β -diversity accounted for by the environment was statistically greater than either the spatial or shared components as hypothesized (Table 4). The average environmental component of β -diversity was greater than the spatial or shared components in 59% of the 753 analyses. However, the differences between the average percentage of the environmental, spatial and shared components was small (Table 4). That is, the environment was less than 10% greater than either the spatial or shared components in 47% of the 442 analyses where the environment was statistically greater than the spatial or shared components. Also, the percentage of β -diversity attributed to spatial and shared components exceeded the environmental component in 30% and 28% of the 753 analyses, respectively. The percentage of unexplained variation in β -diversity did not differ between studies where the environmental

component was greater than the spatial or shared components versus the studies where the environment was less than the spatial or shared components ($F_{1, 751} = 1.16; p = 0.28$).

Hypothesis 2: Simple versus Complex analyses

Contrary to Hypothesis 2, *complex* analyses did not account for more variation in β -diversity than *simple* analyses (Figure 4). MANOVA showed no difference between the *simple* and *complex* groups (Pillai = 0.01, $F_{4, 748} = 2.0, p = 0.09$) in the percentage of variation in β -diversity accounted for by the correlated components of VP analyses (E, S, E+S and unexplained variation). Also, none of the tests on the individual components of β -diversity were significant (Table 4). The percentage of the total variation in β -diversity explained by factors in the *complex* category ranged from 0.0% to 32.5% with an average of 4.6%. The number of environmental descriptors in *simple* analyses ranged from 3 to 60 with an average of 13.9, whereas the number of environmental descriptors per analysis in *complex* studies ranged from 1 to 40 with an average of 17.4.

Hypothesis 3: taxonomic identifications versus functional groups

Evidence supporting Hypothesis 3 was non-conclusive and was likely effected by a small number of analyses using functional groups. MANOVA showed at least one significant difference between analyses using taxonomic identifications versus functional groups (Pillai = 0.04, $F_{4, 747} = 5.2, p = <0.001$) among the dependent variables (E, S, E+S and unexplained variation). However, there was no difference between taxonomic identifications and functional groups when the components of VP analyses were analyzed separately (Table 5 and Figure 5). The ANOVA tests likely lacked sufficient degrees of freedom to detect a significant increase in the percentage of β -diversity accounted for by functional groups compared to taxonomic identifications.

Discussion

One of the primary purposes of VP analyses is to show the relative importance of metacommunity models in determining patterns of species composition across sites measured as β -diversity (Algarte *et al.* 2014). That objective should remain inconclusive if most of the variation in β -diversity remains unexplained. The perception that environmental filtering is a more common process than neutral dynamics associated with regional dispersal limitations in determining variation in β -diversity is frequently expressed in the literature (e.g. Van der Gucht *et al.* 2007; Algarte *et al.* 2014), and was partially supported by this review. That is, E accounted for a greater percentage of β -diversity than either S or E+S in 59% of the 753 VP analyses, whereas S accounted for more variation in β -diversity than E in only 30% of the VP analyses. However, the strength of the relationship between the environment and β -diversity was weak. The environment only accounted for an average of 14.5% of β -diversity, and there were only 7 out of 753 VP analyses where environmental filtering accounted for $\geq 50\%$ of β -diversity. Similarly, S and E+S only accounted for an average of 9.5% and 8.3% of β -diversity, respectively, and both accounted for $\geq 50\%$ of β -diversity in only 3 out of 753 VP analyses. Also, none of the components of β -diversity (E, S, E+S and unexplained) based on functional groups differed from VP analyses based on species, and the amount of unexplained variation in functional groups (61%) was similar to species groups (65.6%).

The most striking result of this review is that, on average, 65.5% of the variation in β -diversity remained unexplained, and the unexplained variation was $\geq 50\%$ in 616 out of 753 VP analyses (82%). Common explanations for a high percentage of unexplained variation include: the omission of important factors, and limitations of VP analyses (Gilbert & Bennett 2010b). We will also outline the potential importance of stochasticity.

Omission of important factors

This is a compelling explanation because it seems likely that we could omit important factors related to patterns of β -diversity because of the complexity of nature. However, investigators will not knowingly omit potentially important factors. Thus, we might expect studies with more environmental factors to account for more variation in β -diversity because they have a greater probability of including important unknown factors. We found no relationship between the number of environmental descriptors and the percentage of unexplained variation based on 595 VP analyses, which suggests that the number of environmental descriptors is less important than the type of environmental descriptor.

At the beginning of VP analyses, 98% of the environmental descriptors were abiotic factors operating at the scale of local communities (Cottenie 2005; Soininen 2014; Soininen 2016). In recent years more complex environmental descriptors have been added with the expectation of decreasing the percentage of unexplained variation in β -diversity (Yamaura *et al.* 2008; Gavilanez & Stevens 2013). We found no difference in the percentage of unexplained variation, the percentage of E, or the percentage of E+S between studies based on local abiotic descriptors versus more complex descriptors. This suggests that investigators may be unaware of the factors that determine patterns of β -diversity. There are almost an endless number of ways to represent the complexities of nature and we may unknowingly omit important factors. For example, there are at many different ways to represent the potential effects of temperature (e.g. daily and annual averages and standard deviations, seasonal rates of change, number of degree days, etc.) on the distribution and abundance of organisms in streams (Poff & Ward 1992; Maheu *et al.* 2016). Also, we appear to have a limited ability to represent some complex factors in VP analysis that are known to determine community structure in some ecosystems.

For example, no studies have successfully incorporated the complexity of indirect effects (trophic cascades, keystone species, apparent competition) into VP analyses despite their known importance in determining the distribution and abundance of species (e.g. Menge 1995).

The argument that unmeasured and/or unknown factors account for the high percentage of unexplained variation in β -diversity comes with an important caveat. That is, we will always be able to claim that unknowable or unmeasured factors might alter our perception of reality. This is one of the basic philosophical principles of science, that “the world of nature is indeterminate. The behavior of the particle is uncertain and therefore the behavior of the atom is an uncertainty.” (Schneer 1960). Thus, we could never reject the validity of VP analyses or the metacommunity perspective unless we have enough data to weigh the balance of evidence. The weight of evidence from 753 analyses indicates that we rarely will account for a high percentage of the total variation in β -diversity using VP analyses.

Limitations of VP

Although VP analyses have their limitations, there is no inherent statistical barrier preventing VP from accounting for most of the variation in β -diversity. Simulations have shown that VP analyses can, to a limited extent, underestimate E and overestimate S (Gilbert & Bennett 2010a), and spatial factors may not unambiguously distinguish E from S (Smith & Lundholm 2010). Since there is no optimal formula for VP analyses, it is likely that some of the unexplained variation in β -diversity can be attributed to the multiple ways of executing VP procedures.

Using taxonomic or functional data based on a single snap-shot in time is potentially a fatal weakness of most all VP analyses (Beaudrot & Marshall 2011; Gavilanez & Stevens 2013). That is, only 19 VP studies collected samples in multiple seasons (e.g. Roussel *et al.* 2010;

Alahuhta *et al.* 2013), only 39 in multiples years (e.g. Mykra *et al.* 2007; Blundo *et al.* 2012) and only 10 studies collected samples in multiple seasons for multiple years (e.g. Potapova & Charles 2002; Chen *et al.* 2015). Thus, 10 out of 170 of the VP studies up to 2020, 6%, used snap-shot data to infer the relative importance of metacommunity dynamics.

Snap-shot data assumes that the processes that produce patterns of diversity do not vary over time. This assumption is clearly false as all processes of community assembly generate patterns of diversity over time, and the relative importance of processes can vary in time and space (Eros *et al.* 2012; Erős *et al.* 2014; Tonkin *et al.* 2016). Many metacommunities (maybe all metacommunities) fluctuate between periods of high and low connectivity on seasonal, annual, and longer temporal scales (Zhao *et al.* 2017). Thus, the importance of metacommunity dynamics also fluctuate because they depend on levels of dispersal and connectivity. That is, the proportion of species showing mass effects (high connectivity), environmental sorting (intermediate connectivity) and neutral dynamics (limited connectivity) should vary across time within a metacommunity. Also, temporal fluctuations of connectivity can vary spatially within a metacommunity and can fluctuate according to predictable cycles or unpredictably as a function of environmental variation (e.g. disturbances).

We may underestimate the importance of environmental sorting or source-sink dynamics using snap-shot data because of temporal variation in habitat heterogeneity. Local communities can select for a different suite of species at different times as environmental conditions change (Arrieira *et al.* 2017), and local communities can fluctuate between being a source or a sink (Mouquet & Loreau 2003). Thus, the temporal scale of sampling needs to match temporal fluctuations of connectivity and habitat heterogeneity if we hope to detect metacommunity dynamics using descriptive data, like β -diversity. Also, it is unlikely that we could ever use

snap-shot data to detect patch-dynamics where inferior competitors frequently go extinct in local communities but persist in the metacommunity by rapid colonization of “open” sites. How can a snap-shot of the species composition in local communities determine if the absence of a species was caused by competitors (patch-dynamics) or by random extinction (neutral dynamic)?

Stochasticity

Part of the unexplained variation in β -diversity may be attributed to the effects of stochastic forces on community assembly. The effects of chance related to dispersal dynamics will initially determine the suite of species colonizing localities in a newly formed metacommunity. Thus, the species composition (β -diversity) of localities may be poorly correlated with environmental conditions. According to the environmental filtering hypothesis, if the colonization rate of new species across the metacommunity is high, species sorting can quickly produce the suite of species from the regional pool of species best adapted to environmental conditions. Consequently, many species will be correlated with the variation in environmental conditions and the percentage of unexplained variation in β -diversity will decline.

The efficiency of sorting will decline if environmental conditions or the regional pool of species change faster than the rate of colonization. The isolation of local communities will slow the rate of colonization, such that rates of sorting may not keep pace with the influx of new species to the region as the ranges of species expand and contract, and as new species are generated that require time to disperse. Priority effects may also leave a lasting imprint on community composition largely determined by chance associated with dispersal if the suite of initial colonists can monopolize resources and resist the invasion of other species ((Vannette & Fukami 2014; Fukami 2015). Also, environmental conditions in local communities may change faster than rates of colonization and sorting such that local communities never converge on a

suite of “best adapted” species that are highly correlated with environmental descriptors (Mutz *et al.* 2017).

The distribution of many species in a metacommunity may be random because rare populations are susceptible to extinction by demographic and environmental stochasticity. (Preston 1962) showed that most species in a local community are rare, and the abundances of species can fluctuate over time at a locality and among different localities within a metacommunity (Shimadzu *et al.* 2013). Thus, the distribution of many species in a metacommunity may be random because populations comprised of few individuals are susceptible to extinction by demographic and environmental stochasticity (Lande 1993; Engen *et al.* 1998). Chance can have a greater effect on the distribution of rare species because rare species produce little to no propagule pressure to overcome demographic and environmental stochasticity. The classic “blinking lights” metapopulation model describes species that randomly go extinct in some patches while other patches contain populations capable of colonizing or re-colonizing “empty” patches (Levins 1969). Which populations go extinct and which patches are colonized is determined by chance (Richter-Dyn & Goel 1972; Rajakaruna *et al.* 2013). This stochastic model may apply to the many rare species in a metacommunity.

Environmental filtering assumes that all species can disperse to all localities in a metacommunity. All else being equal, rare species should have a lower probability of dispersing to all suitable localities in a metacommunity than abundant species because rare species produce fewer individuals available for dispersal. If dispersal is random, then the distribution of rare species across an array of local communities will be random. Dispersal limitations and random extinctions are properties of neutral theory (Hubbell 2001).

If the availability of essential resources varies among the localities of a metacommunity, then species sorting will be strong among specialist species. That is specialists will predictably occur at sites with their essential resources. However, the sorting process will be weakened by an abundance of generalist species that can tolerate a broad range of environmental conditions and use a variety of resources which allows them to persist in most localities of metacommunity. Consequently, the distribution generalists will be poorly correlated with environmental descriptors. In one of the few studies to test this hypothesis, Pandit et al. (2005) showed that specialists responded primarily to environmental factors, while generalists responded primarily to spatial factors.

Although null models have played a valuable role in distinguishing deterministic patterns and random variation (Gotelli), they have not been incorporated into VP analyses of metacommunity dynamics. For example, the patch dynamics model assumes that extinctions are attributed to strong hierarchical competition where inferior competitors are continually on the move colonizing open localities as they are driven extinct at other localities (Pickett & White 1985). This should create a “checker board” spatial pattern where inferior competitors are absent at locations occupied by superior competitors (Cody *et al.* 1975). A null model could help to distinguish random extinctions, which are part of neutral dynamics, from expected spatial patterns derived from metacommunity processes.

Characteristics of outlier studies

The percentage of unexplained variation in 7 out of a total of 170 VP studies was <16%. Why did these studies account for most of the variation in β -diversity? First, efficient dispersal and rapid population growth in microbial species (e.g. bacteria, diatoms, etc). The implication is that microbes can rapidly disperse to all localities in a metacommunity where even slight

environmental differences can select for the suite of “species” best adapted to local conditions. Rapid sorting is attributed to a high turnover rate of species from the species pool and rapid population growth rates that over-ride short-lived priority effects (e.g. Potapova & Charles 2002; Van der Gucht *et al.* 2007; Almasia *et al.* 2016). Second, a set of species with similar traits apparently influenced by a single strong selective force, such as interspecific competition leading to habitat partitioning in forest primates ((Beaudrot & Marshall 2011). Third, studies conducted at the scale of entire continents can average over deterministic and stochastic variation at smaller scales to reveal patterns at large scales (Ahl & Allen 1996) in seaweeds (Smit *et al.* 2017) and temperate forests (Xing & He 2019). Fourth, dispersal limitations associated with a strong decay of community similarity with the distance between suitable habitats (E+S) in highly structured environments with very different habitat types (Nakagawa 2014).

Conclusions

In most cases, VP procedures are too simplistic, and any inferences about the relative importance of metacommunity models must be tentative if based on snap-shot data where the vast majority of variation remains unexplained. A single snap shot will not detect changes in spatial patterns caused by changes in environmental conditions. Similarly, levels of connectivity can vary over time leading to different spatial patterns and thus, conclusions about the relative importance of metacommunity dynamics.

Studies that accounted for the majority of variation in β -diversity suggest that VP analyses might best apply to a small set of species with similar traits. Although the best suite of traits may vary among groups of species, resource requirements, population growth rates, and dispersal capabilities appear to provide valuable information (Tonkin *et al.* 2016).

Variance partitioning as a method to account for variation in β -diversity has its own shortcomings due to the inconsistency in the manner with which it is applied across studies. Three techniques for variance partitioning make up the bulk of the literature, those being partial redundancy analysis (pRDA) (Tonkin 2016), the most common method, followed by partial canonical correlation analysis (CCA) (Potapova *et al* 2002), and partial Mantel tests (Zhang *et al* 2012). The way in which authors represent the spatial component (e.g. distance matrices, PCNM, MEM) also varies between articles and has been shown to cause variation in the amount of unexplained variation from the same data in a single study (Gilbert & Bennett 2010). It is recommended that authors familiarize themselves with the methods and potential shortcomings of variance partitioning as discussed in Legendre & Legendre 1998, Bennet & Gilbert 2010, and Bauman *et al* 2019 before using variance partitioning methods to analyze their ecological data.

Metacommunity models are also too simplistic partly because they fail to account for random variation generated by stochastic forces. All forces that effect the distribution and abundance of species have both deterministic and stochastic components (Denny & Gaines 2000; Hubbell 2001; Lande *et al.* 2003; Schindler 2019). Null models applied to metacommunity analyses may provide valuable insight on the relative importance of stochastic forces and deterministic processes that shape the structure of metacommunities. Perhaps if we build from the bottom-up using small sets of species and null models we will eventually arrive at a consensus on the relative importance of metacommunity dynamics.

References

1. Abdelaal, M. (2017). Current status of the floristic composition in Wadi Hagul, Northwest Suez Gulf, Egypt. *Rendiconti Lincei-Scienze Fisiche E Naturali*, 28, 81-92.
- 2.

- Ahl, V. & Allen, T.H. (1996). Hierarchy Theory: A Vision, Vocabulary, and Epistemology.
3. Alahuhta, J., Kanninen, A., Hellsten, S., Vuori, K.-M., Kuoppala, M. & Hamalainen, H. (2013). Environmental and spatial correlates of community composition, richness and status of boreal lake macrophytes. *Ecological Indicators*, 32, 172-181.
 4. Algarte, V.M., Rodrigues, L., Landeiro, V.L., Siqueira, T. & Bini, L.M. (2014). Variance partitioning of deconstructed periphyton communities: does the use of biological traits matter? *Hydrobiologia*, 722, 279-290.
 5. Allen, P.J. & Bennett, K. (2008). *SPSS for the health & behavioural sciences*. Thomson, Australia.
 6. Almasia, R., Caru, M., Handford, M. & Orlando, J. (2016). Environmental conditions shape soil bacterial community structure in a fragmented landscape. *Soil Biology & Biochemistry*, 103, 39-45.
 7. Arrieira, R.L., Schwind, L.T.F., Bonecker, C.C. & Lansac-Tôha, F.A. (2017). Environmental factors exert predominant effects on testate amoeba metacommunities during droughts in floodplains. *Austral Ecology*, 42, 210-217.
 8. Ateş, C., Kaymaz, Ö., Kale, H.E. & Tekindal, M.A. (2019). Comparison of Test Statistics of Nonnormal and Unbalanced Samples for Multivariate Analysis of Variance in terms of Type-I Error Rates. *Computational and Mathematical Methods in Medicine*, 2019, 2173638.
 9. Bauman, D., Vleminckx, J., Hardy, O.J. & Drouet, T. (2019). Testing and interpreting the shared space-environment fraction in variation partitioning analyses of ecological data. *Oikos*, 128, 274-285.
 10. Beaudrot, L.H. & Marshall, A.J. (2011). Primate communities are structured more by dispersal limitation than by niches. *Journal of Animal Ecology*, 80, 332-341.
 11. Blundo, C., Malizia, L.R., Blake, J.G. & Brown, A.D. (2012). Tree species distribution in Andean forests: influence of regional and local factors. *Journal of Tropical Ecology*, 28, 83-95.
 12. Carvalho, R.A. & Tejerina-Garro, F.L. (2015). Environmental and spatial processes: what controls the functional structure of fish assemblages in tropical rivers and headwater streams? *Ecology of Freshwater Fish*, 24, 317-328.
 13. Chen, S., Slik, J.W.F., Mao, L., Zhang, J., Sa, R., Zhou, K. *et al.* (2015). Spatial patterns and environmental correlates of bryophyte richness: sampling effort matters. *Biodiversity and Conservation*, 24, 593-607.
 - 14.

- Cody, M.L., MacArthur, R.H., Diamond, J.M., Diamond, P.G.J. & Cody, P.B.M.L. (1975). *Ecology and Evolution of Communities*. Belknap Press of Harvard University Press.
15.
Cottenie, K. (2005). Integrating environmental and spatial processes in ecological community dynamics. *Journal of Ecology*, 93, 1175-1182.
16.
Denny, M. & Gaines, S. (2000). *Chance in Biology Using Probability to Explore Nature*. Princeton University Press.
17.
Engen, S., Bakke, Ø. & Islam, A. (1998). Demographic and Environmental Stochasticity- Concepts and Definitions. *Biometrics*, 54, 840-846.
18.
Erős, T., Sály, P., Takács, P., Higgins, C.L., Bíró, P. & Schmera, D. (2014). Quantifying temporal variability in the metacommunity structure of stream fishes: the influence of non-native species and environmental drivers. *Hydrobiologia*, 722, 31-43.
19.
Eros, T., Saly, P., Takacs, P., Specziar, A. & Biro, P. (2012). Temporal variability in the spatial and environmental determinants of functional metacommunity organization - stream fish in a human-modified landscape. *Freshwater Biology*, 57, 1914-1928.
20.
Fukami, T. (2015). Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annual Review of Ecology, Evolution, and Systematics*, 46, 1-23.
21.
Gavilanez, M.M. & Stevens, R.D. (2013). Role of environmental, historical and spatial processes in the structure of Neotropical primate communities: contrasting taxonomic and phylogenetic perspectives. *Global Ecology and Biogeography*, 22, 607-619.
22.
Gianuca, A.T., Engelen, J., Brans, K.I., Hanashiro, F.T.T., Vanhamel, M., van den Berg, E.M. *et al.* (2018). Taxonomic, functional and phylogenetic metacommunity ecology of cladoceran zooplankton along urbanization gradients. *Ecography*, 41, 183-194.
23.
Gilbert, B. & Bennett, J.R. (2010a). Partitioning variation in ecological communities: do the numbers add up? *Journal of Applied Ecology*, 47, 1071-1082.
24.
Gilbert, B. & Bennett, J.R. (2010b). Partitioning variation in ecological communities: do the numbers add up? , 47, 1071-1082.
25.
Guichard, F. (2017). Recent advances in metacommunities and meta-ecosystem theories. *Frontiers in Ecology and Evolution*, 6, 610-610.
26.
Holyoak, M., Leibold, M.A., Holt, R.D. & Meeting, E.S.o.A. (2005). *Metacommunities: Spatial Dynamics and Ecological Communities*. University of Chicago Press.
27.
Holyoak, M. & Loreau, M. (2006). Reconciling empirical ecology with neutral community models. *Ecology*, 87, 1370-1377.

28. Hubbell, S.P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.
29. Hutchinson, G.E. (1951). Copepodology for the Onithologist. *Ecology*, 32, 571-577.
30. Jablonski, D. (2017). Approaches to Macroevolution: 2. Sorting of Variation, Some Overarching Issues, and General Conclusions. *Evolutionary biology*, 44, 451-475.
31. Kadmon, R. & Tielbörger, K. (1999). Testing for Source-Sink Population Dynamics: An Experimental Approach Exemplified with Desert Annuals. *Oikos*, 86, 417-429.
32. Lande, R. (1993). Risks of Population Extinction from Demographic and Environmental Stochasticity and Random Catastrophes. *The American naturalist*, 142, 911-927.
33. Lande, R., Engen, S. & Saether, B.-E. (2003). Stochastic Population Dynamics in Ecology and Conservation.
34. Landeiro, V.L., Bini, L.M., Melo, A.S., Oliveira Pes, A.M. & Magnusson, W.E. (2012). The roles of dispersal limitation and environmental conditions in controlling caddisfly (Trichoptera) assemblages. *Freshwater Biology*, 57, 1554-1564.
35. Landsman, K. (2016). On the notion of free will in the Free Will Theorem.
36. Legendre, P. & Legendre, L.F.J. (1998). *Numerical Ecology*. Elsevier Science.
37. Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F. *et al.* (2004). The metacommunity concept: a framework for multi-scale community ecology. 7, 601-613.
38. Maheu, A., Poff, N.L. & St-Hilaire, A. (2016). A Classification of Stream Water Temperature Regimes in the Conterminous USA. *River Research and Applications*, 32, 896-906.
39. Menge, B.A. (1995). Indirect Effects in Marine Rocky Intertidal Interaction Webs: Patterns and Importance. *Ecological Monographs*, 65, 21-74.
40. Miller, R.G. (1981). *Simultaneous Statistical Inference*. Springer-Verlag.
41. Mouquet, N. & Loreau, M. (2003). Community Patterns in Source-Sink Metacommunities. *American Naturalist*, 162, 544-557.
42. Mutz, J., Underwood, N. & Inouye, B.D. (2017). Time since disturbance affects colonization dynamics in a metapopulation. *Journal of Animal Ecology*, 86, 1065-1073.
- 43.

- Mykra, H., Heino, J. & Muotka, T. (2007). Scale-related patterns in the spatial and environmental components of stream macroinvertebrate assemblage variation. *Global Ecology and Biogeography*, 16, 149-159.
44. Nakagawa, H. (2014). Contribution of environmental and spatial factors to the structure of stream fish assemblages at different spatial scales. *Ecology of Freshwater Fish*, 23, 208-223.
45. Pickett, S.T.A. & White, P.S. (1985). *The Ecology of Natural Disturbance and Patch Dynamics*. Elsevier Science.
46. Pielou, E.C. (1972). Niche Width and Niche Overlap: A Method for Measuring Them. 53, 687-692.
47. Poff, N.L. & Ward, J.V. (1992). Heterogeneous Currents and Algal Resources Mediate in situ Foraging Activity of a Mobile Stream Grazer. *Oikos*, 65, 465-478.
48. Potapova, M.G. & Charles, D.F. (2002). Benthic diatoms in USA rivers: distributions along spatial and environmental gradients. *Journal of Biogeography*, 29, 167-187.
49. Preston, F.W. (1962). The Canonical Distribution of Commonness and Rarity: Part I. *Ecology*, 43, 185-215.
50. R Development Core Team (2019). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria.
51. Rajakaruna, H., Potapov, A. & Lewis, M. (2013). Impact of stochasticity in immigration and reintroduction on colonizing and extirpating populations. *Theoretical Population Biology*, 85, 38-48.
52. Richter-Dyn, N. & Goel, N.S. (1972). On the extinction of a colonizing species. *Theor Popul Biol*, 3, 406-433.
53. Rosenberg, A. & McShea, D.W. (2008). *Philosophy of Biology: A Contemporary Introduction*. Routledge.
54. Roussel, E., Crec'hriou, R., Lenfant, P., Mader, J. & Planes, S. (2010). Relative influences of space, time and environment on coastal ichthyoplankton assemblages along a temperate rocky shore. *Journal of Plankton Research*, 32, 1443-1457.
55. Schindler, D.E. (2019). The phenology of migration in an unpredictable world. *Journal of Animal Ecology*, 88, 8-10.
56. Schneer, C.J. (1960). *The Evolution of Physical Science: The Development of the Major Ideas in the Physical Sciences from the Earliest Times to the Present*. Grove Press.
- 57.

- Shimadzu, H., Dornelas, M., Henderson, P.A. & Magurran, A.E. (2013). Diversity is maintained by seasonal variation in species abundance. *BMC Biol*, 11, 98-98.
- 58.
- Smit, A.J., Bolton, J.J. & Anderson, R.J. (2017). Seaweeds in Two Oceans: Beta-Diversity. *Frontiers in Marine Science*, 4, 404.
- 59.
- Smith, J.M. (1989). The Causes of Extinction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 325, 241-252.
- 60.
- Smith, T.W. & Lundholm, J.T. (2010). Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography*, 33, 648-655.
- 61.
- Soininen, J. (2007). The distance decay of similarity in boreal stream communities. *Ecological Society of America Annual Meeting Abstracts*.
- 62.
- Soininen, J. (2014). A quantitative analysis of species sorting across organisms and ecosystems. 95, 3284-3292.
- 63.
- Soininen, J. (2016). Spatial structure in ecological communities – a quantitative analysis. 125, 160-166.
- 64.
- Tonkin, J.D., Stoll, S., Jaehnig, S.C. & Haase, P. (2016). Contrasting metacommunity structure and beta diversity in an aquatic-floodplain system. *Oikos*, 125, 686-697.
- 65.
- Van der Gucht, K., Cottenie, K., Muylaert, K., Vloemans, N., Cousin, S., Declerck, S. *et al.* (2007). The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 20404-20409.
- 66.
- Vannette, R.L. & Fukami, T. (2014). Historical contingency in species interactions: towards niche-based predictions. 17, 115-124.
- 67.
- Vasudevan, A. (2018). Chance, determinism and the classical theory of probability. *Studies in History and Philosophy of Science Part A*, 67, 32-43.
- 68.
- Vellend, M. (2016). *The Theory of Ecological Communities (MPB-57)*. Princeton University Press.
- 69.
- Whittaker, R.H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, 30, 279-338.
- 70.
- Wilson, D.S. (1992). Complex Interactions in Metacommunities, with Implications for Biodiversity and Higher Levels of Selection. *Ecology*, 73, 1984-2000.
- 71.
- Winegardner, A.K., Jones, B.K., Ng, I.S., Siqueira, T. & Cottenie, K. (2012). The terminology of metacommunity ecology. *Trends Ecol Evol*, 27, 253-254.

72.

Xing, D. & He, F. (2019). Environmental filtering explains a U-shape latitudinal pattern in regional β -diversity for eastern North American trees. *Ecology Letters*, 22, 284-291.

73.

Yamaura, Y., Katoh, K. & Takahashi, T. (2008). Effects of stand, landscape, and spatial variables on bird communities in larch plantations and deciduous forests in central Japan. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, 38, 1223-1243.

74.

Zhao, K., Song, K., Pan, Y., Wang, L., Da, L. & Wang, Q. (2017). Metacommunity structure of zooplankton in river networks: Roles of environmental and spatial factors. *Ecological Indicators*, 73, 96-104.

Tables and Figures

Table 1. Environmental descriptors used in *complex* studies. N1 and N2 are the total number of articles and the total number of data sets in each *complex* category, respectively. Note that number of data sets (167) is greater than the number of articles (50) because many articles contained data sets that applied to multiple categories.

Complex categories	N1	N2
Species Interactions	10	50
Human Effects	8	22
Topography	3	4
Climate and Regional factors	21	63
History	7	24
Phylogenetic Relatedness	1	1

Table 2. The number of articles and variance partitioning data sets used to calculate the average percentage of environmental, spatial, shared, and unexplained variation in β -diversity by Cottenie (2005), Soininen (2014 and 2016), and in this study. A dash indicates averages that were not calculated.

Authors	Articles	Data Sets	Environmental	Spatial	Shared	Unexplained
Cottenie	66	158	22.4%	15.4%	11.1%	50.0%
Soininen	54	326	26.1%	11.0%	-	-
Lamb and Rader	170	753	14.5%	9.5%	8.3%	65.6%

Table 3. Analyses showing percentages for environmental (E), spatial (S), and shared (E+S) components of β -diversity that were outliers based on VP studies prior to 2020. Outliers for each component are in bold lettering. Some articles contained more than one data set, and some data sets occurred in multiple components. Also, the component percentages in some data sets do not sum to 100% because factors that are not shown in the table (e.g. phylogeny) accounted for part of the total variation in β -diversity.

Taxa	E	S	E+S	Unexpl	Citation
Soil bacteria	76.0	12.5	0.0	11.5	Almasia et al., 2016
Stream macroinvertebrates	45.3	0.0	0.0	54.7	Astorga et al., 2011
Temperate forest ants	44.1	23.4	0.0	32.5	Liu et al., 2018
"	44.3	23.5	0.0	32.2	"
"	69.9	11.9	0.0	18.2	"
"	68.8	10.5	0.0	20.7	"
River diatoms	45.1	15.9	11.9	27.2	Potapova & Charles, 2002
Floodplain spiders	51.5	44.2	0.0	39.7	Tonkin et al., 2016
Lake Bacteria	52.0	19.0	18.0	11.0	Van Der Gucht et al., 2007
Rock pool macroinvertebrates	46.3	8.4	2.2	43.2	Vanschoenwinkel et al., 2007
"	47.1	9.1	2.1	41.6	"
Temperate forest trees	68.0	0.0	29.0	3.0	Xing et al., 2019
"	68.5	0.0	27.5	4.0	"
River zooplankton	47.9	0.0	0.0	52.3	Zhao et al., 2017
Stream macroinvertebrates	4.0	54.0	24.0	18.0	Astorga et al., 2011
Tropical forest primates	13.0	61.0	21.0	5.0	Beaudrot & Marshall, 2011
Soil mites	2.0	40.0	2.0	56.0	Caruso et al., 2012
Stream fish	12.9	34.9	17.5	34.7	Carvalho & Tejerina-Garro, 2015
Mosses and Liverworts	7.0	31.0	16.0	46.0	Chen et al., 2015
Tropical broad-leaved forest	2.9	34.8	27.8	34.5	Legendre et al., 2009
Stream fish	3.0	41.0	38.0	18.0	Leprieur et al., 2009
"	6.0	43.0	0.0	51.0	"
Thermophilous vegetation	13.1	32.7	1.6	52.6	Szymura & Szymura, 2013
Floodplain invertebrates	24.1	33.0	0.0	63.4	Tonkin et al., 2016
"	8.8	40.7	0.0	69.6	"
"	22.8	60.8	0.0	35.0	"
Floodplain spiders	51.5	44.2	0.0	39.7	"
Ground-dwelling arthropods	0.0	40.0	24.0	36.0	Guo et al., 2019

Alpine plants	16.1	19.0	37.4	27.7	Anthelme et al., 2003
Oribatid mites	13.7	12.2	31.0	43.0	Borcard & Legendre, 1994
Bryophytes	23.0	8.0	29.0	40.0	Chen et al., 2015
Diatoms	3.0	22.0	40.0	35.0	Dong et al., 2016
Forest trees	22.0	7.0	32.0	39.0	Gazol & Ibanez 2010

Table 3. continued

Taxa	E	S	E+S	Unexpl	Citation
Lake shoreline plants	5.0	22.1	35.2	37.7	Girdler et al., 2008
Diatoms	5.8	13.8	30.0	50.4	Heino & Soininen, 2007
Microbes	11.0	3.6	28.0	57.1	Heino et al., 2014
Understory vegetation	35.5	4.3	31.5	28.8	Hu et al., 2013
Oribatid Mites	9.5	19.5	29.0	42.0	Ingimarsdottir et al., 2012
terrestrial ferns	0.6	14.6	27.8	56.9	Jones et al., 2011
Broad-leaved evergreen trees	2.9	34.8	27.8	34.5	Legendre et al., 2009.
Freshwater fish	3.0	41.0	38.0	18.0	Leprieur et al., 2009
Alpine plants	15.5	8.0	29.1	47.4	Li et al., 2011
Vascular Plants	36.4	6.5	28.5	28.6	Lobo et al., 2001
River Fish	5.2	12.6	68.2	14.0	Nakagawa, 2014
Ichthyoplankton	12.0	1.4	31.6	50.5	Roussel et al., 2010
Seaweeds	1.8	13.7	82.4	2.1	Smit et al., 2017
River Fish	6.0	2.0	30.0	35.0	Koster et al., 2007
Tropical climbing plants	8.0	19.0	28.0	28.0	Thonhofer et al., 2015
Floodplain invertebrates	6.8	0.0	29.9	64.0	Tonkin et al., 2016
Aquatic macroinvertebrates	1.1	12.3	40.8	25.5	Zhang et al., 2019
Soil bacteria	76.0	12.5	0.0	11.5	Almasia et al., 2016
Tropical forest primates	13.0	61.0	21.0	5.0	Beaudrot & Marshall, 2011
River fish	5.2	12.6	68.2	14.0	Nakagawa, 2014
River diatoms	36.5	25.9	21.3	16.3	Potapova & Charles, 2002
Seaweeds	1.8	13.7	82.4	2.1	Smit et al., 2017
Lake bacteria	52.0	19.0	18.0	11.0	Van Der Gucht et al., 2007
Temperate forest trees	68.0	0.0	29.0	3.0	Xing & He, 2019
"	68.5	0.0	27.5	4.0	"

Table 4. Tukey-Kramer post-hoc pairwise comparisons between the mean percentage of the total β -diversity attributed to environmental, spatial, and shared components in variance partitioning analyses.

Hypothesis	Difference Between Means	<i>p</i> -value
Env > Spatial	5.1%	<0.001
Env > Shared	6.2%	<0.001
Spatial > Shared	1.2%	0.06

Table 5. ANOVA (R Development Core Team 2019) between *simple* (n=600) versus *complex* (n=153) data sets using transformed values ($\ln X+1$) of the percentage of the total variation in β -diversity attributed to environmental, shared, and unexplained components of variance partitioning analyses. *Simple* studies included only local abiotic descriptors. *Complex* studies included local abiotic descriptors, plus, complex environmental descriptors shown in Table 1. A Bonferroni correction set the level of significance at $p = 0.0125$.

Dependent Variable	Hypothesis	F-statistic	<i>p</i> -value
Environmental	Complex > Simple	5.7 _{1, 751}	0.02
Shared	Complex > Simple	2.7 _{1, 751}	0.10
Unexplained	Simple > Complex	2.5 _{1, 751}	0.12

Table 6. ANOVA (R Core Team 2019) between studies using taxonomic identifications (n = 726 data sets) versus functional groups (n = 26 data sets) to calculate the percentage of the total variation of β -diversity attributed to environmental, spatial, shared, and unexplained components of variance partitioning analyses. A Bonferroni correction set the level of significance at $p = 0.0125$.

Dependent variable	Hypotheses	F-values	<i>p</i> -values
Environmental	Functional > Taxonomic	1.5 _{1, 750}	0.6
Spatial	Functional > Taxonomic	1.4 _{1, 750}	0.2
Shared	Functional > Taxonomic	0.4 _{1, 750}	0.5
Unexplained	Taxonomic > Functional	5.3 _{1, 750}	0.1

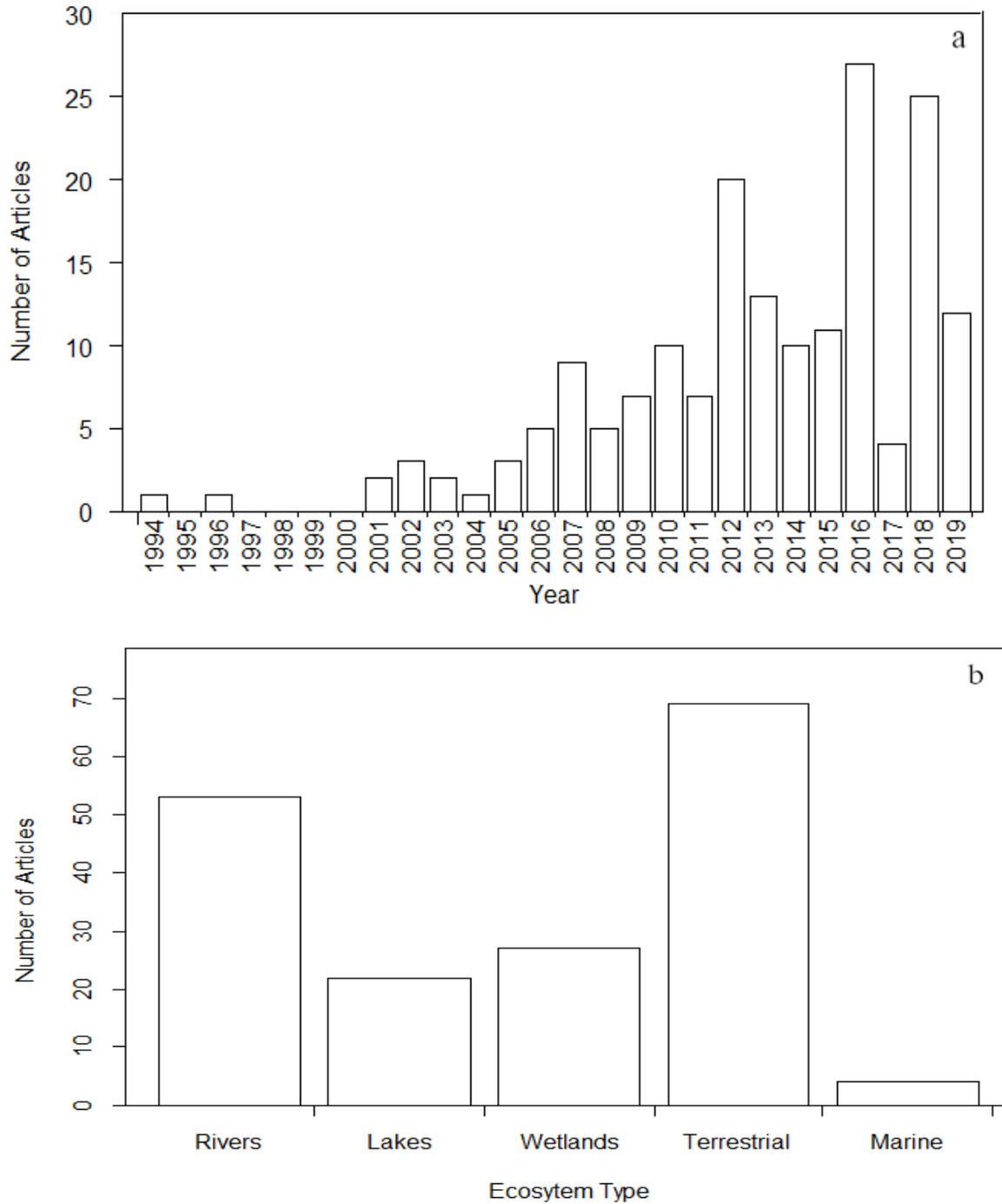


Figure 1. Number of peer-reviewed articles that partitioned the variation in β -diversity into 4 additive components (environmental, spatial, shared and unexplained) per year (a) and by ecosystem type (b).

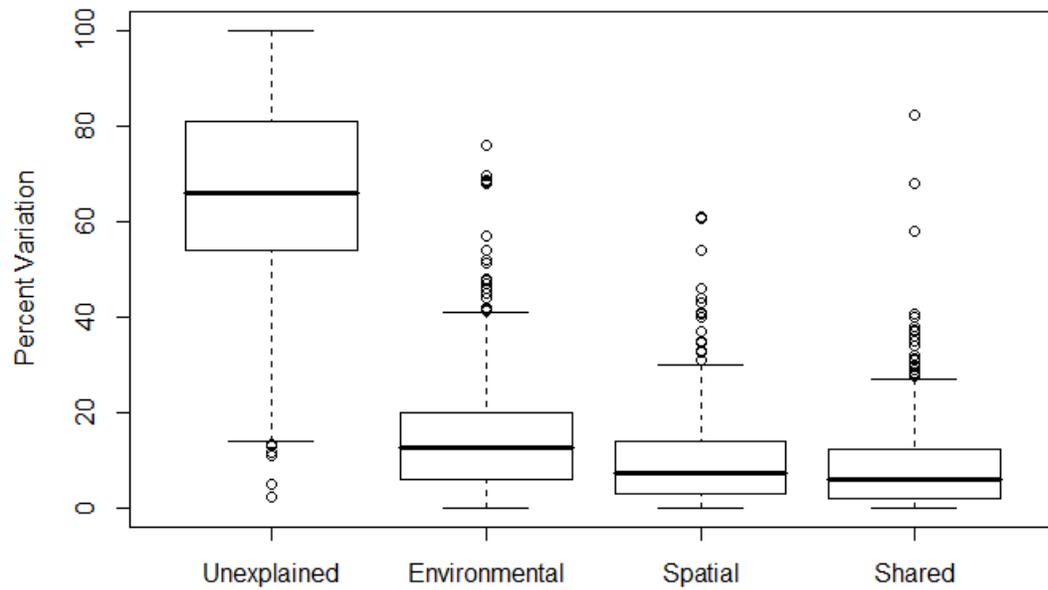


Figure 2. Box plots of data sets collected prior to 2020 showing the variation around the average unexplained percentage of β -diversity, and the average percentage of β -diversity attributed to the environment, spatial factors, and the shared components of variance partitioning analyses. Open circles show outliers as data sets that fell outside the 95% confidence limits. This summary is based on 170 articles and 753 data sets.

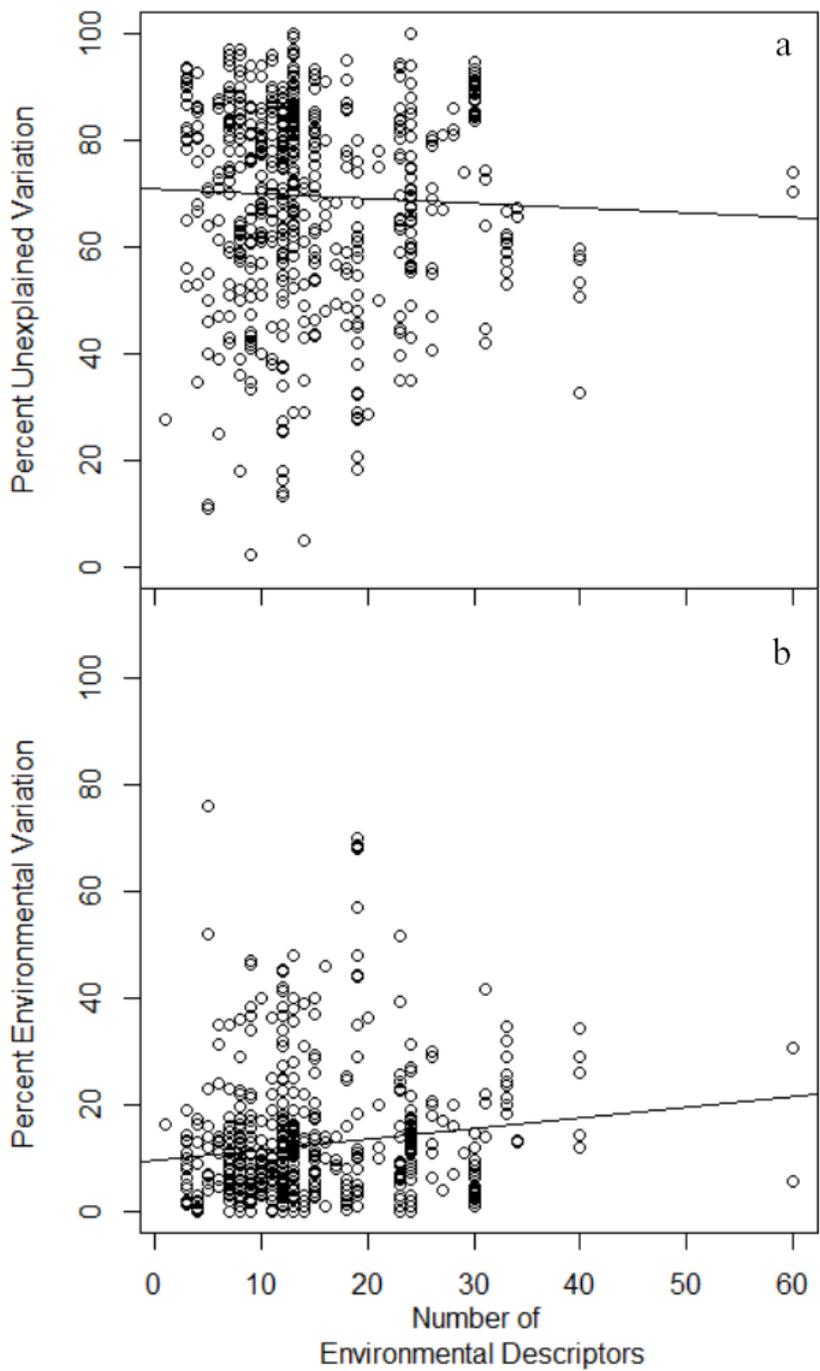


Figure 3. Scatter plots showing the relationship between the number of environmental descriptors and the percentage of unexplained variation (a) and environmental variation (b) based on 595 site-by-taxa data sets used to partition β -diverity.

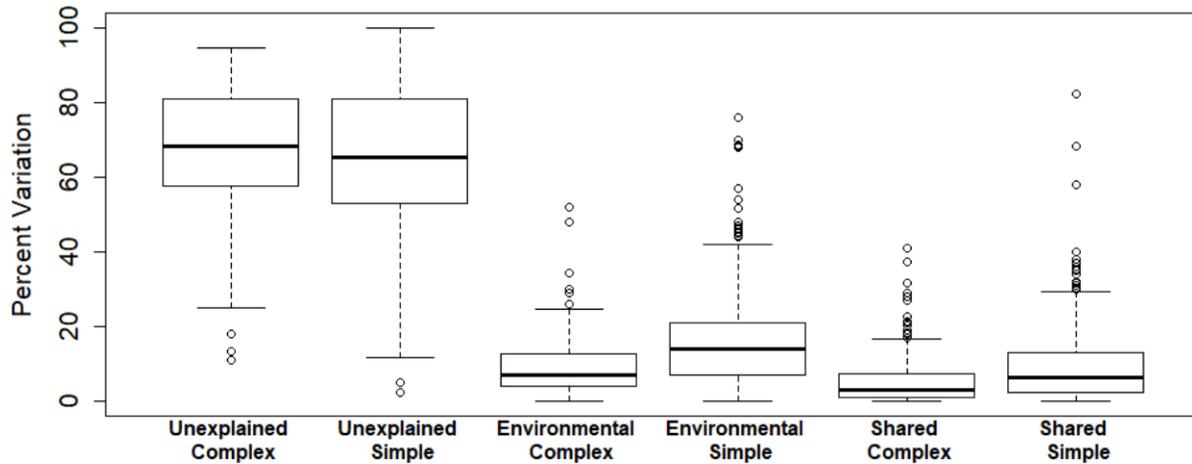


Figure 4. Box plots showing the percentage of the total variation in β -diversity attributed to unexplained, environmental, and shared components of variance partitioning analyses between studies using *simple* versus *complex* environmental descriptors. *Complex* studies included environmental descriptors other than local abiotic factors, whereas, *simple* studies included only local abiotic factors.

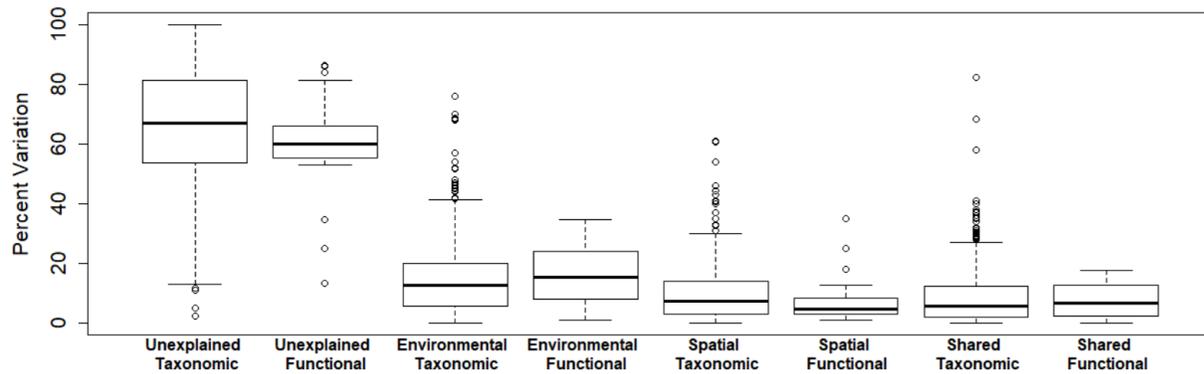


Figure 5. Box plots showing the percentage of the total variation in β -diversity attributed to unexplained, environmental, spatial, and shared components of variance partitioning analyses comparing studies using taxonomic diversity versus functional diversity.