Perspective

Is local biodiversity declining or not? A summary of the debate over analysis of species richness time trends

Bradley J. Cardinale, Andrew Gonzalez, Ginger R.H. Allington, Michel Loreau

A R T I C L E   I N F O

Keywords:
Species extinction
Invasion
Biotic homogenization
Alpha
Beta
Gamma diversity

A B S T R A C T

Recently, a debate has developed over how biodiversity is changing across the planet. While most researchers agree species extinctions are increasing globally due to human activity, some now argue that species richness at local scales is not declining as many biologists have claimed. This argument stems from recent syntheses of time-series data that suggest species richness is decreasing in some locations, increasing in others, but not changing on average. Critics of these syntheses (like us) have argued there are serious limitations of existing time-series datasets and their analyses that preclude meaningful conclusions about local biodiversity change. Specifically, authors of these syntheses have failed to account for several primary drivers of biodiversity change, have relied on data poor time-series that lack baselines needed to detect change, and have unreasonably extrapolated conclusions. Here we summarize the history of this debate, as well as key papers and exchanges that have helped clarify new issues and ideas. To resolve the debate, we suggest future researchers be more clear about the hypotheses of biodiversity change being tested, focus less on amassing large datasets, and more on amassing high-quality datasets that provide unambiguous tests of the hypotheses. Researchers should also keep track of the contributions that native versus non-native species make to biodiversity time trends, as these have different implications for conservation. Lastly, we suggest researchers be aware of pros and cons of using different types of data (e.g., time-series, spatial comparisons), taking care to resolve divergent results among sources to allow broader conclusions about biodiversity change.

1. Introduction

Over the past few years, a scientific debate has developed over how biodiversity is changing across the planet. Most researchers agree that species extinctions at the global scale are occurring much faster than what is ‘normal’ in the fossil record (Barnosky et al., 2011). The majority of researchers would also agree that biodiversity is generally declining at most locations across the planet, especially in areas that have experienced direct human impact. This view is, in fact, sufficiently ingrained in the minds of biologists that select disciplines (e.g., Conservation Biology) and fields of study (e.g., Biodiversity and Ecosystem Functioning) often take local species extinctions as a given, and a primary motivation for their work. But a group of ecologists has recently begun to claim they have amassed a body of evidence showing that species richness is, in fact, not declining at local spatial scales across the globe, and that the objectives of conservation need to be re-examined (Dornelas et al., 2014; Hillebrand et al., 2018; McGill et al., 2015; Vellend et al., 2013).

The claim that local diversity is not in decline stems primarily from analysis of time-series data of biological monitoring programs. While the data themselves are not controversial, the analyses of the data and conclusions that have followed have been controversial, which has led to a series of exchanges between proponents and critics of the use of time-series to quantify local diversity change (Cardinale, 2014; Dornelas et al., 2014; Gonzalez et al., 2016; Hillebrand et al., 2018; McGill et al., 2015; Vellend et al., 2013; Vellend et al., 2017b). But the exchanges have taken place at an assortment of working groups and meetings, and been dispersed across a variety of journals (mostly ecological). The goal of this paper is to provide practitioners of biodiversity conservation with some background on the debate, summarize the key
papers and exchanges to date that help clarify new ideas, and then offer our perspective on how to move forward towards a resolution. A successful resolution is critically important to the field of conservation biology, as well as other fields of study, where the effectiveness of management decisions depends on our ability to accurately predict how biodiversity is changing at various scales across the planet.

2. Background

The foundation for the current debate over biodiversity change traces back to papers addressing the countervailing effects of human induced species loss and invasion on local biodiversity ( McKinney and Lockwood, 1999; Olden and Poff, 2003; Sax and Gaines, 2003; Sax et al., 2002). Sax and Gaines (2003) were among the first to show that rates of species introductions into large regions (e.g., islands, mainland ecosystems) sometimes equal or exceed rates of species extinctions. When introductions equal or exceed extinctions, species richness can remain constant or even increase, rather than decline as has often been presumed by those who cite the negative impacts of invasive species on biodiversity ( Clavero and García-Bethoux, 2005; Molnar et al., 2008; Wilcove et al., 1998). If species gains outpace extinctions in ecosystems, and we accept that species extinctions are exceeding rates of speciation at the global scale, then the relationship between species diversity at different scales of observation ( gamma (γ) = alpha (α) × beta (β) ) suggests that beta-diversity - the turnover of species among locations on the planet – is generally in decline ( Whitaker, 1960). The resulting prediction is that the world’s biota is being homogenized as non-indigenous and locally expanding species replace local biota ( Olden et al., 2004).

Since publications by Sax and the ensuing discussion about biotic homogenization, it has been increasingly argued that loss of beta-diversity deserves more attention by practitioners and managers in biodiversity conservation ( Gering et al., 2003; Magurran, 2016; Olden, 2006). But even as attention has turned towards better quantification of changes in local diversity ( alpha (α) ) and species turnover ( beta (β) ), it has become clear that we generally lack the types of datasets that are needed to broadly assess alpha- and beta-diversity for the average location on Earth. Indeed, several authors have emphasized the paucity of long-term monitoring programs that assess a broad range of organisms across the terrestrial land surface, and oceans of the planet ( Duarte et al., 1992; Green et al., 2005; Henry et al., 2008; Pereira and Cooper, 2006; Sheil, 2001). Of the programs that do exist, most have been uncoordinated, non-uniform in methods and coverage, and are not easily accessible by scientists attempting to perform data syntheses.

In an effort to overcome these limitations, several studies began collating the data needed to quantify local changes in biodiversity around the globe, and to determine the extent to which communities are changing. Though these studies and working groups share a common goal, they have taken different approaches. Some have focused on estimating diversity change using spatial comparisons in which measures of species diversity in reference sites are compared to measures of diversity in habitats that have been modified by human activities. This was the approach taken by meta-analyses that made spatial comparisons between disturbed and undisturbed reference sites ( e.g., Alroy, 2017; Aronson et al., 2014; Benayas et al., 2009; Gerstner et al., 2014; Moreno-Mateos et al., 2017; Murphy and Romanuk, 2014; Newbold et al., 2016; Newbold et al., 2015).

In contrast to efforts that have used spatial comparisons, a second group of data syntheses has taken a different approach in which researchers have collated time-series data from studies that have made repeated measurements of biodiversity at individual locations around the planet ( Dornelas et al., 2014; Elahi et al., 2015; Hillebrand et al., 2018; Vellend et al., 2013). Studies focused on analysis of time-series datasets have generally come to a different conclusion. While these analyses have confirmed extensive turnover in species composition across locations, and across time at single locations, they have not found evidence of systematic declines in local species richness. Rather, these data syntheses have shown that species richness in some locations has increased through time, whereas in other locations it has declined. These opposing trends have been offsetting such that there has been no net change in local species richness at the average location represented in the datasets.

Because of the surprising results and their publication in prominent journals, data syntheses by those like Vellend et al. (2013) and Dornelas et al. (2014) received considerable attention in the popular media. For example, a write-up of the Dornelas et al. (2014) paper in Science World Report stated: “As our climate changes, species are disappearing—or that’s what’s commonly assumed. Now, though, it looks like that might not be the case. Scientists have re-examined data from 100 long-term monitoring studies done around the world and have found that the number of species hasn’t changed much, or has actually increased over time ( Griffin, 2014).” Authors of the original syntheses wrote in follow-up papers their analyses had overturned the long-standing view that species richness at local scales across the globe is declining, contrary to what many ecologists and conservation biologists have claimed ( McGill et al., 2015; Vellend 2017) took this message to the public in a subsequent OpEd in American Scientist magazine, writing: “It is unsettling to have one’s view of the world called into question—in this case I had to face evidence that is contrary to the conventional wisdom in conservation biology imparted to me in the 1990s. Biodiversity is not generally declining at all spatial scales: Declines at the global scale are not generally seen at the regional scale and occur only in particular scenarios at the local scale.”

Despite the claim that historical views on biodiversity loss have been overturned, this claim has been controversial. Several critiques and criticisms of the Vellend et al. (2013) and Dornelas et al. (2014) data syntheses have been published ( Cardinale, 2014; Gonzalez et al., 2016; Isbell et al., 2015), and working groups organized at the Integrative Biodiversity Research Center in Germany, the Biodiversity Research Center in Canada, and the Quebec Center for Biodiversity Science in Canada have brought together the original authors and their critics to debate key issues in person. Unfortunately, this exchange has occurred in scattered venues and publications, making the debate hard to follow for those who have not been directly involved in the exchange.

We believe the outcome of the debate over local biodiversity change is critically important to the field of conservation biology. If, as analyses of time-series datasets suggest ( Dornelas et al., 2014; McGill et al., 2015; Vellend et al., 2017a; Vellend et al., 2013), species richness is not generally in decline at local scales as has long been presumed, the historical tools used for biodiversity conservation may need to be revised, and the trends reported in many textbooks need to be rewritten. If, however, critics are correct about the limitations of conclusions drawn from time-series data, then it may be premature to suggest that historical views about biodiversity loss have been overturned. In the remainder of this paper, we summarize the key papers and exchanges that have helped clarify new issues and ideas, after which, we offer some suggestions on how to move towards a resolution.

3. Summary of the debate

In this section we summarize key arguments from four primary papers that exemplify the current debate ( Dornelas et al., 2014;
Table 1
Summary of key characteristics of the Vellend et al. (2013) and Dornelas et al. (2014) data syntheses of time-series that are the foundation of debate over local species change.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Vellend et al. (2013)</th>
<th>Dornelas et al. (2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of studies summarized</td>
<td>346</td>
<td>100</td>
</tr>
<tr>
<td>Habitats represented</td>
<td>Terrestrial</td>
<td>Mostly marine, with a few freshwater and terrestrial sites</td>
</tr>
<tr>
<td>Organismal focus</td>
<td>Vascular plants</td>
<td>Multiple: mammals, birds, fishes, invertebrates, and plants</td>
</tr>
<tr>
<td>Number of locations</td>
<td>&gt; 16,000</td>
<td>430,324</td>
</tr>
<tr>
<td>Time-scales</td>
<td>5 to 261 years, median = 20</td>
<td>3 to 51 years, median = 13</td>
</tr>
<tr>
<td>Spatial-scales</td>
<td>0.04 to 13,000 m²</td>
<td>0.01 to 2,080,000 m²</td>
</tr>
<tr>
<td>Response variable</td>
<td>log response ratios (LRR) used to quantify proportional change in richness from initial to final year</td>
<td>Slopes from regressions of species richness as a function of time</td>
</tr>
<tr>
<td>Finding</td>
<td>Some LRR &gt; 0, some &lt; 0. On average, LRR were = 0</td>
<td>Some slopes &gt; 0, some &lt; 0. On average, slopes were = 0</td>
</tr>
<tr>
<td>Conclusion</td>
<td>On average, there has been no net loss of local species richness, as ecologists and conservation biologists have historically claimed.</td>
<td></td>
</tr>
</tbody>
</table>

Gonzalez et al., 2016; Vellend et al., 2013; Vellend et al., 2017b). Vellend et al. (2013) and Dornelas et al. (2014) represent the initial data syntheses of time-series that are the foundation of the debate. While additional analyses of time-series have since been published (e.g., Elahi et al., 2015; Hillebrand et al., 2018), the two original papers are still the most widely cited and influential analyses of biological monitoring programs. Gonzalez et al. (2016) is the most comprehensive critique of the use of time-series data to measure biodiversity change, and incorporated concerns that had been voiced by others (Cardinale, 2014; Isbell et al., 2015). Vellend et al. (2017b) was written by many authors of the original papers (Dornelas et al., 2014; Vellend et al., 2013), and represents the latest response to concerns about their two original papers.

Table 1 provides a summary of the key characteristics of the Vellend et al. (2013) and Dornelas et al. (2014) data syntheses, which we have complemented with a brief description of the methods of the two papers in the Supplemental information, along with their finding of no net change in local species richness. After publication of these two papers, Gonzalez et al. (2016) published a critique, and argued their conclusion of no net change in local species richness was questionable on two grounds: (1) the time-series datasets collated for these syntheses were not spatially representative of where biodiversity is located around the globe, nor was it representative of the primary drivers of local biodiversity change, and (2) the datasets used by the two syntheses were mostly composed of data-poor time-series that lacked historical baselines, making it impossible to accurately characterize biodiversity change. Gonzalez et al. (2016) went on to criticize the authors for overextrapolating their conclusions to global scales and to other fields of study. We summarize each of these arguments below, as well as Vellend et al.’s (2017b) response in order to clarify differences in perspective that have emerged, and to emphasize areas that are in need of more research.

3.1. Spatial representation of biodiversity and drivers of change

While both Vellend et al. (2013) and Dornelas et al. (2014) claimed to have completed systematic meta-analyses representing global trends in biodiversity, Gonzalez et al. (2016) showed such claims were inaccurate. After completing a formal spatial analysis of the studies that were included in the two syntheses, Gonzalez et al. (2016) showed those datasets laid 8 to 41 standard deviations outside of globally representative samples of where biodiversity is located across Earth’s land surface and oceans, and 12 to 32 standard deviations outside of globally representative samples of where humans have most impacted Earth’s land surface and oceans (see Fig. 1 in Gonzalez et al., 2016). Geographic biases are common in data syntheses, including most that have attempted summarized local biodiversity change. However, most authors of data syntheses constrain their conclusions, and do not claim their results to be globally representative of diversity change, as Vellend et al. and Dornelas et al. did. Therefore, the problem is not that Vellend et al. and Dornelas et al. used geographically biased datasets; rather, the problem is that these authors over-extended their statements and conclusions beyond what could be reasonably concluded from their data. The claims that their results were globally representative of what is happening to biodiversity across the planet were demonstrably false.

Gonzalez et al. (2016) also argued that Vellend et al. (2013) and Dornelas et al. (2014) misled readers in suggesting their syntheses had accounted for the dominant human drivers of biodiversity change. For example, Vellend et al. (2013) claimed their synthesis represented plant species diversity change “from all major vegetation types, including areas under profound and direct human influence.” Despite this claim, the synthesis did not include any studies focused on the most prominent drivers of terrestrial diversity loss, such as the conversion of terrestrial habitats into agricultural or urban systems (Fig. 1A-B). Instead, Vellend et al.’s (2013) dataset was composed of studies that had been designed for a mixture of purposes, including those documenting:

- biodiversity change caused by variety of types of natural disturbances (Fig. 1C), such as volcanic eruptions, altered fire regimes, altered grazing, and more;
- recovery of biodiversity from natural disturbances (Fig. 1D), such as successional studies of plants recovering from volcanic eruptions, fires, grazing, and more;
- biodiversity change in response to a variety of types of anthropogenic perturbations (Fig. 1E), such as logging, pollution, cattle grazing, invasive species, and more;
- recovery of biodiversity from anthropogenic disturbances (Fig. 1F), such as from logging, pollution, cattle grazing, invasive species, and more;
- biodiversity change in remnant patches of ecosystems that had not undergone direct habitat loss or conversion (Fig. 1G).

Gonzalez et al. (2016) showed that some of these types of studies were more represented than others. In particular, studies focused on the recovery of historically disturbed patches (e.g., forest succession after logging) were substantially over-represented in the Vellend et al. (2013) synthesis, whereas studies in areas undergoing habitat loss (e.g., deforestation) were under-represented (see Fig. 2 in Gonzalez et al., 2016). The dataset was, in fact, heavily biased towards ecosystems in North America and Europe that were in various stages of recovery from historic disturbances like logging (e.g., Fig. 1F). While geographic bias was acknowledged in the original paper, the authors did not control for this bias in their analyses, and their conclusions were extrapolated well beyond the locations represented in the dataset. Furthermore, Vellend et al. (2013) did not control for the over-representation of studies focused on recovery of historically disturbed patches.

Similarly, Gonzalez et al. (2016) noted that it was difficult to interpret what the collection of studies collated by Dornelas et al. (2014) represented since the authors did not report which drivers of biodiversity change were being studied in the monitoring programs they summarized. The spatial analyses completed by Gonzalez et al. (2016) showed the datasets were mostly composed of time-series taken from coastlines along the eastern U.S., coastlines along European countries and, to a lesser extent the Antarctic shelf. Because some of these areas
B.J. Cardinale et al.

Fig. 1. Examples of land use change that can impact biodiversity differently. Time 1 shows an intact forest. By time 2, parts of the forest have been converted to urban (A) or agricultural habitat (B). Other parts have experienced natural (C) or anthropogenic (E) disturbances, or are recovering from disturbance (D, F). Some of the original forest remains as fragmented patches (G) where biodiversity may still be influenced by surrounding habitat loss, invasive species, climate change, or other indirect effects related to habitat destruction. Because the various land use types are expected to influence biodiversity differently, one cannot collate all of these land use types into a single data-synthesis and accurately quantify biodiversity change without accounting for the heterogeneity in drivers. The only way to accurately estimate change would be to quantify diversity change in each land-use type, and then perform an analysis that is weighted by the areal extent of land-use types in the landscape. In contrast to this approach, data syntheses like that by Vellend et al. (2013) only considered C-G.

(i.e. coasts of the U.S. and Europe) tend to be some of the most heavily impacted oceanic systems according to maps of human influence generated by marine researchers (Halpern et al., 2008), the Dornelas et al. (2014) synthesis may have over-represented areas of human impact on the oceans relative to a globally representative sample. Gonzalez et al. (2016) argued that it is impossible to know whether this bias over-represents human impacts that are presently ongoing (e.g., overharvesting), or over-represents habitats that are now recovering from historic impacts (e.g., improvement management of fisheries).

In their most recent paper, Vellend et al. (2017b) acknowledged the geographic biases in their original data syntheses. But they went on to argue that spatial biases are not likely to influence their conclusions about no net loss of local species richness because (1) the few datasets they have from under-represented regions of the globe (South America, Asia, Australia, and Africa) generally showed increases rather than decreases in species richness, (2) over-representation of human influence on the oceans in the Dornelas et al. (2014) synthesis suggests that human activities do not, as a generality, produce declines in local species richness, and (3) the time-series data they have collated represent the best empirical evidence available for documenting local diversity change. For point (1), the authors assumed the limited set of studies they collated from South America, Asia, Australia, and Africa are broadly representative of those parts of the globe, which is questionable given Gonzalez et al.'s (2016) analyses showing that areas experiencing habitat loss were significantly under-represented in the data syntheses. In making point (2), Vellend et al. (2017b) did not address Gonzalez et al.'s (2016) warning that, because the authors did not report the drivers of diversity change being studied, it is impossible to know if the over-representation of human impacts in the Dornelas et al. (2014) synthesis was caused by over-representation of sites that are currently experiencing human impacts (e.g., overharvesting), or that are presently recovering from human impacts. Point 3 sidestepped Gonzalez et al.'s (2016) warning that just because data exist, that does not mean those data are appropriate for the question at hand. Indeed, if the best data presently available are not representative of biodiversity change – either spatially across the globe, or of the primary drivers of change – then those data are simply not appropriate for addressing the original question about how local species richness is changing across the planet.

In their most recent paper, Vellend et al. (2017b) further argued that criticisms about biases in their datasets (both spatial, and in representation of drivers of diversity change) are “typical of those that could be directed at any ecological meta-analysis.” Aside from the fact that repeating others’ mistakes is not a valid scientific defense, this comment does not reflect how meta-analyses are typically used. Historically, meta-analyses have been used to evaluate the efficacy of some set of experiments that share a common purpose and design, and conclusions from meta-analyses have traditionally been constrained to drawing inferences about the results of that particular set of experiments. Vellend et al. (2013) and Dornelas et al. (2014) used meta-analyses in a fundamentally different way. Rather than summarizing the outcomes of experiments that shared a common purpose and design, Vellend et al. (2013) and Dornelas et al. (2014) collated time trends from observational studies that focused on a heterogeneous array of ecological drivers, many of which would be expected, a priori, to influence biodiversity in profoundly different ways (much like the heterogeneity in Fig. 1).

Experts in the use of meta-analyses have repeatedly warned about the misinterpretation of data that can result when authors do not control for heterogeneity among studies with differing purposes and designs, and fail to control for differences in the a priori expectations of studies being summarized (Hillebrand and Cardinale, 2010; Koricheva and Gurevitch, 2014; Nakagawa et al., 2017; Ridolo et al., 2017; Whittaker, 2010). To illustrate the misinterpretations that can arise, consider the scenario in Box 1 where a researcher performs a meta-analysis to determine how diet impacts body weight in laboratory mice.

In this example, a researcher comes to incorrect conclusions due to a misattribution to the results of that particular set of experiments. Vellend et al.’s (2013) data synthesis in which the authors took studies that focused on a heterogeneous array of ecological drivers, many of which would be expected, a priori, to influence biodiversity in profoundly different ways. This is the same logical
Box 1
Logical flaws in data syntheses.

Hypothetical meta-analysis. Imagine a hypothetical data synthesis in which a medical researcher interested in how diet impacts body weight collates two different types of studies. One set of studies is composed of experiments that fed mice a high calorie diet to quantify weight gain. The second set of studies is composed of experiments where mice have been put on a low calorie diet to quantify weight loss. The researcher pools these two types of experiments into a single dataset, after which, s/he performs a meta-analysis to quantify that change in mouse body weight as a function of diet. The researcher concludes some mice gain weight, others lose weight. But on average, there is no net change in body weight of mice. Consequently, the researcher concludes that diet has no directional impact on mouse body weight.

Logical flaw 1. Clearly, the medical researcher’s logic and conclusions are flawed. The flaw stems from the fact that the researcher collated a heterogeneous set of studies that were designed for entirely different purposes and then subjected them to a common analysis without properly accounting for the heterogeneity. This is the same logical flaw that Vellend et al. (2013), and likely Dornelas et al. (2014), made in their meta-analyses where they collate studies that documented changes in biodiversity following perturbations (fires, grazing, and other disturbances) with studies that documented the recovery of biodiversity as ecosystems were recovering from the exact same perturbations. This may help explain why the authors found no net change in biodiversity across studies.

Logical flaw 2. Imagine that after completing their meta-analysis, the medical researcher goes on to conclude that because there is no net change in mouse weight in response to diet, diet must not influence a third, unmeasured variable – mouse diabetes. Clearly, this conclusion would also be flawed because the researcher extrapolated his/her results to a third, unmeasured variable. This is the same logical flaw committed by Vellend et al. (2013) who, after concluding from their meta-analysis that net biodiversity is not changing at local scales, went on to “question the widespread use of ecosystem function experiments to argue for the importance of biodiversity conservation in nature.”

because the authors did not report the drivers of diversity change. Gonzalez et al. (2016) re-analyzed Vellend et al.’s (2013) dataset to account for differences in a priori expectations, and showed that species richness significantly declined through time when analyses were limited to communities that were responding to disturbances. However, those declines were offset by significant recovery of species richness after cessation of those disturbances, which caused the net effects to center on zero (become non-significant). Gonzalez et al. (2016) argued that, based on the same logic in Box 1, it is not appropriate to take a study that has documented a decline in local richness after an impact, collate that with a study documenting recovery after the impact has ceased, and then analyze the two trends together and conclude there has been no net change in local biodiversity. Gonzalez et al. (2016) was concerned that doing this risks producing inaccurate conclusions that humans have had no net impact on local richness. In their latest response, Vellend et al. (2017b) disagreed, writing “In a world with fire, grazing, logging, and other disturbances of varying intensity and frequency, to include only the effects of such disturbances and to ignore locations recovering from past disturbances constitutes the introduction of a bias in itself.” It seems to us that all involved in this debate agree that any proper accounting of biodiversity change across the globe must account for areas that have not only been impacted by perturbations, but that are in various phases of recovery following perturbations. Failing to sample time series from disturbed and recovering habitats in proportion to their global representation of habitat is sure to generate misleading conclusions about biodiversity change. Still, there remains a difference of opinion on whether studies documenting responses of local diversity to multiple stressors should be collated to produce a single global estimate of net change. This issue still needs to be resolved, and we provide some recommendations at the end of this paper.

3.2. Baselines and the quality of time-series data

Gonzalez et al. (2016) expressed two concerns about the quality of data used by Vellend et al. (2013) and Dornelas et al. (2014), and whether those data were adequate to test hypotheses about changes in local species richness. One concern was that the datasets lacked any historical baselines or reference conditions. Establishment of baselines or reference conditions is routine practice for researchers who work with time-series datasets. Consider, for example, data used by climate change scientists to track global temperatures or ice cover (Fig. 2). A common feature of nearly all data analyses is that researchers present temperature or ice-cover data as ‘anomalies’ that represent deviations relative to some established baseline. This baseline may represent what was normal in the pre-industrial era, or may represent some long-term running average. Regardless of what benchmark is used, a baseline is necessary because one cannot know whether data are ‘abnormal’ until you first define what ‘normal’ is.

Historical baselines are widely used by those who study global extinctions (Barnosky et al., 2011; Ceballos et al., 2015), establish goals for local conservation (Papworth et al., 2009), and who work to restore biodiversity to degraded ecosystems (Benayas et al., 2009). Despite the
Fig. 2. A comparison of time-series data used to quantify various aspects of climate change (panels A–C) with three example time-series similar to those used by Vellend et al. (2013) to quantify biodiversity change of terrestrial plants (panel D). Note that data in A–C are all high-resolution time-series that are compared to a historical baseline, which allows researchers to know if data deviate from a ‘normal’ or ‘expected’ value. In addition, researchers look at long-term trends over many years of data because analysis of any small subset of data points could lead to the conclusion that temperatures (or ice cover) are increasing, decreasing, or remaining stable. In contrast, the hypothetical data in D are not compared to any baseline, and inferences are drawn from data-poor time-series involving a select few points. Image A and B are from the National Oceanic and Atmospheric Administration (NOAA). Image C is from the National Snow and Ice Data Center, University of Colorado, Boulder.

precedents, neither Vellend et al. (2013) nor Dornelas et al. (2014) established baselines for the studies included in their data syntheses (see examples, Fig. 2D). As such, Gonzalez et al. (2016) argued that the analyses in these papers were unable to establish whether changes in species richness at any study location represents stochastic variation around ambient levels of diversity, or whether they represent real deviations (anomalies) from normal conditions. In their most recent response, authors of the two original syntheses claimed that the ‘baselines’ in their analyses were simply the first time-points in the data series: “We recognize that all of these analyses were carried out with respect to baselines determined by the beginning of the time series involved. There will be cases where ecosystems have lost or gained biodiversity before these observations began, but at present we cannot assess the frequency of these different scenarios (Vellend et al., 2017b).” This response simply ignored Gonzalez et al.'s (2016) point that a single datum taken from an unknown date in a longer-term time trend is not a reliable benchmark against which to measure change. As is true for all other fields that use time-series, measuring change through time requires the establishment of accurate, historical baselines.

A related criticism by Gonzalez et al. (2016) was that the datasets collated for the two data syntheses were composed mostly of data poor time-series populated by few data points. The median study used by Vellend et al. (2013) had only two estimates of diversity taken 20 years apart (see examples in Fig. 2D). The dataset collated by Dornelas et al. (2014) was a bit more robust with the median time-series containing 13 years of data, but even these time-series are rather scant for measuring local extirpations. Gonzalez et al. (2016) used simulations in which they randomly sampled small subsets of data from longer time-series of hypothetical monitoring programs to show that analyses of data poor time-series do not accurately represent long-term trends in species richness when those long-term trends are unknown a priori. Vellend et al. (2017b) argued this conclusion was based on two errors in the simulations. First, they argued that Gonzalez et al. (2016) failed to “correct” the randomly sampled data subsets by the time period of the simulated monitoring program to estimate diversity loss over the entire time-series. This response ignored Gonzalez et al.'s (2016) point that the long-term trend for biodiversity is not known a priori (as was true in Vellend et al., 2013), which means we don’t know what correction factor to use. If we had perfect knowledge about the entire time-trend, we wouldn’t need to make statistical estimate in the first place!

The second error, Vellend et al. (2017b) argued, was Gonzalez et al.'s (2016) use of bounded time-intervals that produced a ‘mid-domain’ effect. The mid-domain effect occurs when the middle portions of a time-series are over-sampled due to the finite boundaries of the series (beginning and end points), which Vellend et al. (2017b) argued leads to ‘transient’ trends being over-represented in subsamples. Vellend et al. (2017b) did not offer any explanation of why they expect ‘transient’ trends to be systematically biased in a way that over- or under-represents long-term changes in local species richness. Lastly, Vellend et al. (2017b) concluded that conclusions from the empirical dataset depends on assumptions made in the statistical analyses. They showed that outcomes of the analyses depended on whether intercepts for the time-series are, or are not set to zero, and whether data points with large leverage are, or are not included in the analysis. Vellend et al. (2017b) ended their reassessment by saying “The conclusion, based on simulations, that short time series can provide unreliable estimates of a known trend is simply incorrect.”

We think the response by Vellend et al.'s (2017b) brings up some...
legitimate points. It is worth considering how future simulations might be based on better null models that avoid the mid-domain effect. It is also worth being aware of the rather obvious point that the conclusions of statistical analyses depend on the assumptions made and data used—a point that, in our view, serves primarily to emphasize that existing data is not sufficiently robust to provide a clear answer. However, by focusing their response on minor details of the simulations and analyses, Vellend et al.’s (2017b) essentially dodged the more fundamental and important point that Gonzalez et al. (2016) were trying to make. Gonzalez et al.’s (2016) point was that data-poor time-series that have no baselines are not a reliable way to estimate an unknown trend over time. To illustrate conceptually rather than through data simulations, consider again the field of climate science where researchers routinely use time-series data to try and determine if global temperatures or ice cover are changing through time (Fig. 2). Because of year-to-year variation, one could choose any two data points from the full time-series and, depending on which points are selected, conclude the world is getting warmer, colder, or staying constant. Such year-to-year variation is routinely exploited by climate change skeptics who try to mislead the public by analyzing short subsets of data from the longer time-series to show trends that support their points of view (e.g., the recent debate over the warming hiatus). This is why climate scientists do not draw conclusions about long-term trends in temperature records based on short time-series composed of only a few data points. Nor should biologists attempt to draw conclusions about long-term trends in biodiversity change based on short time-series composed of only a few data points.

3.3. Extrapolation to other fields of study

After suggesting that species richness at local scales across the globe is not, in fact, declining as many ecologists and conservation biologists have claimed, Vellend et al. (2013) went on to argue that their findings weaken the rationale for sub-disciplines in conservation biology that are motivated by local biodiversity loss. The authors focused on the field of Biodiversity and Ecosystem Functioning, which is set of mathematical models and experiments that have shown how local species extinctions impact ecological processes that regulate the functioning of ecosystems. Vellend et al. (2013) questioned the relevance of this field for conservation, saying “we find no general tendency for local-scale plant species diversity to decline over the last century, calling into question the widespread use of ecosystem function experiments to argue for the importance of biodiversity conservation in nature.” This claim has been repeated in subsequent papers and opinion articles (Vellend, 2017; Vellend et al., 2017a).

Gonzalez et al. (2016) criticized Vellend et al. (2013) for over-extrapolating their results to other fields of study noting that, even if one accepts that increases in species richness in some locations have offset species declines in others to produce a distribution with no net change, this pattern does not negate the relevance of research from fields of study like Biodiversity and Ecosystem Function for those locations that have experienced declines. Gonzalez et al. (2016) also argued that Vellend et al.’s (2013) extrapolation of their results to the field of Biodiversity and Ecosystem Functioning was based on a logical flaw. To illustrate, return to Box 1 and imagine that after the researcher completes his/her meta-analysis, s/he concludes that because experiments have shown no net change in mouse weight in response to diet, diet must not influence a third variable that has been shown to be related to mouse health (e.g., diabetes). Such a conclusion would be nonsensical, but is the same logic used by Vellend et al. (2013) to conclude that because biodiversity is not changing on average, it must not affect another variable that has previously been related to local biodiversity in other studies (e.g., ecosystem functioning). Vellend et al. (2017b) did not address this criticism in their latest response; yet, they continue to suggest that results from their data syntheses are evidence that select fields of study premised on local species losses are not relevant to conservation (Vellend, 2017; Vellend et al., 2017a).

4. Suggestions for moving towards a resolution

While those who have focused on analyses of time-series assert that species richness at local scales across the globe is not declining as many ecologists and conservation biologists have claimed, we believe this argument is premature, needs to be tempered by a more thorough consideration of the criticisms of the data and analyses, and needs to be resolved with abundant evidence to the contrary (see Supplemental Information for a brief summary). As we navigate this debate towards a conclusion, we offer several suggestions for moving towards a quick, and constructive resolution:

4.1. Make hypotheses clear, and data representative

One of the key problems that has fostered this debate is that studies using time-series data (Dornelas et al., 2014; Vellend et al., 2013) have not been particularly clear about their questions and hypotheses, and have not explicitly considered whether the data being used are appropriate to address their questions. Furthermore, the two studies assessed biodiversity change with data from locations that varied more than nine orders of magnitude in area (Table 1), calling all of it “local” without explicitly accounting for scale effects. The lack of clarity about questions and hypotheses, and the lack of attention to the scale of sampling and analyses, has led to a number of overextended statements about the supposed ‘global’ representation of the results, as well as misleading statements about the types and extent of human drivers of biodiversity change being considered.

To avoid similar confusion in the future, we recommend authors take care to address the following questions in their papers: What is the null hypothesis being tested? Which specific drivers of biodiversity change are being quantified? Over which temporal and spatial scales is biodiversity change being measured, and for which taxonomic groups? Researchers should keep in mind that a priori expectations may differ in each case. Indeed, the null expectation for diversity change of trees in a forest being monitored for 20-years at the scale of a hundred hectares is very different than the null expectation for perennial understory plants being monitored for a few years in 1-m plots. This is because the baseline rates of extinction, colonization and turnover are expected to differ vastly between these scenarios.

4.2. Focus on data quality, not quantity

There is nothing inherently wrong with using time-series data to address questions about local biodiversity change. But it is not appropriate to try and quantify local biodiversity change using data poor time-series that have no historical baselines, and which were collated from a heterogeneous collection of studies that were designed to study many different drivers of diversity change, each with a different a priori expectation. It is important to recognize that a greater abundance of poor-quality data will not help resolve the current controversy, and subjecting ever-expanding datasets to the same types of analyses (as Vellend et al., 2017b) did) will not overcome limitations that are inherent in the current methods. What we need are datasets that have clear baselines that tell us what expected values of biodiversity are, and thus, whether data from monitoring programs are aberrant with respect to background variation. This requires that we are clear about temporal benchmarks for the mean and variance, whether they be pre-impact, pre-industrial, or pre-recovery. In many instances, ecological monitoring programs will not have the data required to establish these benchmarks, and we may need to interact with other fields of study, such as paleobiology to establish the proper expectations. When historical baselines are not available, we will need to use multiple types of data, including BACI designs and spatial references as others have done (Aronson et al., 2014; Murphy and Romanuk, 2014; Newbold et al.,
We also recommend that future researchers avoid collating heterogeneous sets of studies designed for different purposes into a single statistical analysis for a data synthesis. We believe it is important to (a) take greater care to parse out the data from studies into more homogeneous subsets that allow for testing of a clearly stated directional hypothesis (e.g., separating impacts from recovery, or parsing studies focused on different drivers of biodiversity change), or alternatively, (b) explicitly control for the heterogeneity among studies with statistical analyses that use major differences among studies in the datasets as predictor variables. In general, researchers need to focus less on trying to amass large datasets for meta-analyses, and pay more attention to making sure the data collected and used in a meta-analysis are appropriate tests of a clear hypothesis.

4.3. Quantify biotic homogenization, not just changes in local species richness

We also think that part of the debate over the Vellend et al. (2013) and Dornelas et al. (2014) data syntheses stems from their continued speculation that biotic homogenization is the primary driver of their findings; yet, neither study actually quantified homogenization. Dornelas et al. (2014) wrote “One potential driver is that intensification of trade and transport, combined with the rapid increase in invasions of exotic taxa, is leading to the homogenization of species composition at local scales … Our results suggest that local and regional assemblages are experiencing a substitution of their taxa, rather than systematic loss.” But like Vellend et al. (2013), Dornelas et al. (2014) did not keep track of which species were native and non-native in the datasets they collated. As such, their studies cannot tell us if the turnover of species at different time-points is due to the replacement of native species by non-indigenous species, or alternatively, if the beta-diversity documented is driven by the turnover of native species, such as might occur when communities are recovering from disturbances.

To quantify biotic homogenization, we need to know the extent to which non-indigenous species are replacing native biota that are going locally extinct (Olden et al., 2004; Smart et al., 2006). One cannot do this by taking simplistic measures of species richness and converting these into estimates of turnover in space or time. Asserting that species richness has remained constant, but the ecosystem has changed in species composition, is not evidence of homogenization (Hillebrand et al., 2018). For purposes of conservation, future studies that use time-series data need to report how diversity is changing for native and non-native species, and to quantify the fraction of native species that were present at time \( t \) that are being replaced by nonindigenous species at time \( t + n \). It would be even more useful for studies to develop predictive models that tell us which species, and which types (e.g., functional types), are becoming more dominant through time, and which are becoming increasingly rare.

4.4. Recognize the pros and cons of different types of data

It would be naïve to think that any one type of study or data synthesis could be globally representative of the plethora of environmental changes that are altering biodiversity across the planet. While time-series data suffer from problems described in this article, spatial comparison suffer limitations as well, such as the difficulty of finding appropriate reference conditions or making untested assumptions about space-for-time substitution. Therefore, we advocate for bringing different types of information together in analyses that allow observations to be geographically weighted to reflect the heterogeneous contributions of both natural and human drivers of change across regions, land use types, and taxa. To illustrate the type of syntheses we would advocate for, assume we could say the following about the various land use types in Fig. 1 with some degree of certainty:

- Fig. 1A. To date, 4.9-billion ha of natural habitat have been converted to cropland and pastures (MEA, 2005). Data syntheses comparing plant diversity in cropland/pasture to that in reference sites suggest a mean 30% of native species go locally extinct (Newbold et al., 2015).
- Fig. 1B. To date, 0.35-billion hectares of the world’s land surface has been converted into urban habitat (Center for International Earth Science Information Network - CIESIN - Columbia University et al., 2011). Based on data syntheses that have compared plant diversity in urban habitats to that in non-urban reference sites, a mean 75% of native plant species go locally extinct in urban environments (Aronson et al., 2014).
- Fig. 1D. At present, XX-million km² of the world’s forests are in recovery after historic logging. Improved time-series (inspired by BioTIME) have shown that species richness of plants increases by XX% in roughly half of these sites, and decreases by XX% in the other half (note: XX%’s mean we don’t yet have reliable numbers, but we should be able to get them).

... and so on for all land use changes that impact biodiversity (i.e. various categories in Fig. 1).

With information of this kind, it would be possible to produce a spatially representative map showing global changes in biodiversity caused by the summed impact of local environmental change. We could then take 100,000 random lat/lon coordinates from this map and obtain the mean expectation for local biodiversity change at a typical location on Earth. That would tell us whether biodiversity at local scales is generally increasing or decreasing. But more importantly, the map would show us where diversity is increasing/decreasing and why, which would be far more useful for conservation efforts. In principle, a similar analysis could be completed for the world’s surface waters, ocean bottoms, and coastlines.

Analyses like these would represent a major advance over our current state of knowledge, which is to debate the validity and superiority of data syntheses that each focus on very different aspects of biodiversity change. The next round of research on biodiversity change should seek to synthesize the syntheses so that we can truly move towards the globally representative estimates of local biodiversity change.

Acknowledgements

We thank F. Chang, J. Chase, D. Hooper, J. Byrnes, C. Godwin, F. Isbell, M. O’Connor, and P. Thomas for comments on earlier versions of the manuscript. Funding: This work was supported by grants from the U.S. National Science Foundation to B.C. (DEB-1046121, EFRI-1332342), a fellowship from the National Socio-Environmental Synthesis Center to G.A. (NSF #DBI-1052875), an NSERC Discovery grant and Killam Fellowship to A.G., and a TULIP Laboratory of Excellence grant to M.L. (ANR-10-LABX-41).

Appendix A. Supplementary information

Supplementary information to this article can be found online at https://doi.org/10.1016/j.biocon.2017.12.021.

References

Biodiversity Synthesis. World Resources Institute, Washington D.C., pp. 86.


