



Biodiversity vs. paleodiversity measurements: the incommensurability problem

Federica Bocchi¹

Received: 3 March 2022 / Accepted: 22 October 2022 / Published online: 10 November 2022
© Springer Nature B.V. 2022

Abstract

Estimating whether the Earth's biota is in the middle of a crisis relies heavily on comparisons between present and past data about biodiversity or biodiversity surrogates. Although the past is a crucial source of information to assess the severity of the current biodiversity crisis, substantive conceptual and methodological questions remain about how paleodiversity and biodiversity are to be properly compared. I argue that to justify claims of a current biodiversity crisis is harder than it appears. More precisely, I claim that paleodiversity measurements are incommensurable with contemporary measurements, given the different ways that biodiversity is conceptualized and operationalized. I conclude by proposing three possible ways of overcoming this incommensurability problem.

Keywords Biodiversity · Paleodiversity · Measurements · Historical sciences · Epistemology of science · Incommensurability

1 Introduction

Estimating whether the Earth's biota is in the middle of a crisis relies heavily on comparisons between present and past data about biodiversity or biodiversity “surrogates” (Sarkar 2005). It is commonly claimed, for instance, that a massive extinction event (the so-called “sixth mass extinction”) is occurring, evidence for which is found in unprecedented extinction rates—100 to 10,000 times higher than the background extinction rate obtained from fossil evidence (Ceballos et al. 2015). In addition, conservationists investigate the past as a resource

This article belongs to the Topical Collection: EPSA21: Selected papers from the biennial conference in Turin

Guest Editors: A.C. Love, C. Marchionni, M. Redei, J. Williamson

✉ Federica Bocchi
fbocchi@bu.edu

¹ Department of Philosophy, Boston University, 745 Commonwealth Avenue, 02215, MA, Boston, USA

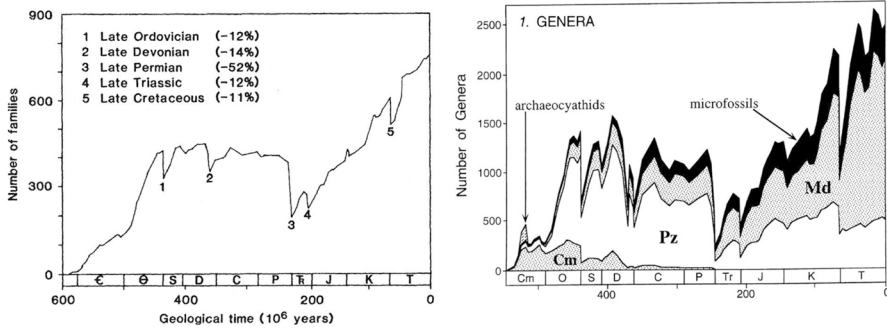


Fig. 1 Sepkoski's paleodiversity curve showing fluctuations of marine invertebrate families, from Raup and Sepkoski (1982). [Reprinted with permission from AAAS]; Sepkoski's (1997) paleocurve at the genus resolution, from Sepkoski (2002)'s data. [Reprinted with permission from Cambridge University Press]

for solving this crisis. The past provides information about environmental tipping points or ecological system responses to perturbations—information which can be put to use in developing conservation and restoration strategies (Benson and Mannion 2012). Although the past is a crucial source of information to assess the severity of the current biodiversity crisis, substantive conceptual and methodological questions remain about how paleodiversity and biodiversity are to be properly compared. This paper will tackle some of these questions.

In Sections 2 and 3, I first present various measurements of biodiversity and of paleodiversity. My intent is to lay the foundation for an answer to the question of whether inferences from paleodiversity measurements to biodiversity estimates are epistemically well-motivated. In Section 4, I argue that justifying such comparative evaluations (e.g. using paleodiversity to show that we are currently facing an unspecified biodiversity crisis) is harder than it appears. More precisely, I claim that paleodiversity measurements are incommensurable with contemporary measurements, given the different ways that biodiversity is conceptualized and operationalized. For example, unlike current biodiversity measures, paleoestimates rely heavily on an understanding of biodiversity as changes in total genus counts, and are mostly based on extrapolations from marine invertebrates' fossils. A paradigmatic example of this is the famous Sepkoski diversity curve of marine diversity fluctuations in deep time, which tracks the timing and tempo of species extinctions from the Cambrian (541 mya) to the present (Fig. 1). Conversely, the understanding of current biodiversity is not reducible to species inventories and loss thereof, as most paleodiversity studies assume. I call this mismatch the “incommensurability problem”. Far from arguing that paleodata are useless in conservation efforts, I emphasize that paleodiversity is not directly commensurable with estimates of contemporary biodiversity *tout court*. I conclude by investigating three possible ways of overcoming this incommensurability problem.

2 Measuring biodiversity

2.1 A disclaimer about biodiversity and paleodiversity measurements

Before discussing the various ways in which biodiversity and paleodiversity are measured and assessed, I must address some terminological and conceptual choices made in this paper. A present or past biodiversity measurement indicates how diverse a system (a soil sample, a clade, an ecosystem) is at time t . This measurement serves to compare the same system over time, or to compare multiple systems.

By “measurement” of biodiversity or paleodiversity I mean a *value* or a *set of values* resulting after a measurement process. I will not investigate the *process* of measuring, namely data collection and correction. A measuring process, for example, usually includes using sampling techniques (like netting, fogging, digging), then data polish and elaboration, computing for systematic errors etc. In this paper, I only consider “measurement outcomes” Tal (2013), a numerical value (\pm uncertainty) expressing a measurable quantity associated with a *measurand*, which in our case is present and past diversity or proxies for them (e.g., taxa richness, phylogenetic history, etc.). Below, I list a few biodiversity indices used by ecologists to measure local and global biodiversity and paleodiversity.

2.2 Species-based biodiversity measurements

In the past few decades, an increasing awareness of anthropogenic impact on Earth has made measuring biodiversity and coping with its loss a hot scientific topic. But biodiversity is a complex concept spanning various levels of biological organization at different spatial and temporal scales, and the question of what exactly ought to be measured is contentious. Despite the proliferation of potentially relevant environmental, taxonomic, and genetic data, there is still a lack of consensus about which parameters best capture biodiversity and biodiversity fluctuations.

To make the concept more empirically tractable in ecology and conservation biology, biodiversity has been operationalized in various ways. Yet, no measurement has been universally validated or considered the most adequate to assess biodiversity outside of specific research fields. This depends, in large part, on how biodiversity is conceptualized in particular research settings and domains. In this section, I review a few indices with respect to which contemporary biodiversity is operationalized and assessed, to support the philosophical claim that different metrics are underpinned by different conceptualizations of biodiversity. The main takeaway from this section is that there currently is no single way to measure biodiversity, because the concept is too complex to be captured by the available quantifications.

SPECIES RICHNESS INDEX. The simplest and most primitive way of measuring biodiversity is by accounting for species *richness*, or species *count*. The species richness index, (S), refers to the absolute number of species in a sample and corresponds to the following equation:

$$S = N,$$

where N is the species count in a sample. 1 is the minimum value of N , and, in general, higher values correspond to more diverse assemblages.¹ The species richness index (S) permits intuitive comparative assessments between distinct datasets. For example, the species richness index answers the question of which dataset obtained from separate ecosystems is more diverse, and the same equation can be used to show biodiversity trends in the same ecosystems across time. However, this measure has a notorious major limitation: it is extremely sensitive to the sampling effect, by which the obtained diversity value potentially changes only in virtue of the collected specimens or the sampled area.² Since the relationship between the number of species and number of sampled individuals is nonlinear—due to what is known as the species accumulation curve—(S) is inadequate when comparing values obtained via unequal samples, which constitute the vast majority of ecological data collection.

MENHINICK SPECIES RICHNESS INDEX. In a classic 1964 paper, ecologist Edward F. Menhinick (1964) suggested a different formula (Menhinick's species' richness index) that partially corrects the sampling bias of S and that would be more useful when comparing samples (of insects) of various sizes. His preferred formula calculates the ratio of the number of species detected in a sample (S) and the squared total of the number of sampled individuals (N):

$$D_{Mn} = \frac{S}{\sqrt{N}}.$$

This index, according to Menhinick, outperforms other fairly common indices (see Gleason 1922; Margalef 1958) when it comes to predicting how species number will change as a function of the sampling effort. This index should therefore be applied to the analysis of uneven field data, if the goal is getting to a more accurate comparative assessment of species richness.

Although an improvement over (S), Menhinick's species' index makes problematic conceptual assumptions (Magurran & McGill 2011). For example, species' relative abundance (how rare or common individual of various species are) is assumed to be normally distributed in the samples, meaning that all species are assumed to be equally represented. Since species' distribution is a critical bit of information about a system's diversity, few conservation biologist or ecologist would validate Menhinick index as adequate measurement of biodiversity.³

Indices of species richness often fail to be validated because they do not accurately capture, among other things, information about how abundant species are

¹ This type of conceptualizing and measuring species richness mostly applies to plant and animal community at a macroscale. Microbial diversity cannot be measured in this way (see O'Malley 2014; Malaterre 2013)

² Suppose you sample a plot p twice, obtaining: a) 100 individuals belonging to 28 species at time $t1$; b) 200 individuals belonging to 36 species at time $t2$. According to S , the second dataset is more diverse. Nevertheless, this conclusion might be an artifact of simply having sampled more individuals.

³ Suppose you have two datasets of 100 individuals, each made up of three species A, B and C, but represented respectively by 30 As, 30 Bs, 40 Cs, and 90 As, 10 Bs, 10 Cs. The diversity value resulting from applying Menhinick's index to the first and second dataset will show two equally diverse communities (dataset one: $D_{Mn} = 3/10$; dataset 2: $D_{Mn} = 3/10$).

(how many individual belong to each species); how abundant species are relative to one another; and species' function or role in a sample.

GINI-SIMPSON DIVERSITY INDEX A popular way of measuring biodiversity understood as species' richness and evenness is the Gini-Simpson index (D_{Gini}), due to Gini (1912) and Simpson (1949), also known as quadratic entropy index. Gini-Simpson's equation indirectly measures the rarity of a species by representing the likelihood of randomly selecting two individuals of the same species from a sample:

$$D_{Gini} = \frac{1}{\sum_{i=1}^s p_i^2},$$

where p refers to the proportion (n/N) of individuals of one of the species (n) divided by the total number of sampled individuals (N) and s is the species number. Despite conveying information about community composition, which is a critical aspect of biodiversity, D_{Gini} presents an occasional extreme non linearity which makes it unsuited to plan conservation efforts in situations of severe species loss.⁴ Logistic limitations are not the only issues in the validation of species-based biodiversity measurements. I now turn to the conceptual aspects.

2.3 Conceptual limitations of richness and abundance measurements

Having highlighted some practical and theoretical shortcomings of a few biodiversity measurements, I here zoom in on a critical conceptual limitation common to both richness indices and richness \times abundance measures (R & A).

As philosophers and scientists have pointed out, species R & A measurements are generally and increasingly understood as either a "component" of biodiversity or as a "sign" of biodiversity (see Sarkar 2005; Maclaurin & Sterelny 2008; Santana 2018), but not as a measurement of biodiversity *itself*. This means that measurements which reduce biodiversity to R & A cannot ultimately be validated as exhaustive measurements because they do not capture all of what is meant by biodiversity.⁵

The awareness that R & A metrics only provide limited information about biodiverse systems has inspired various projects aimed at listing which aspects of biological systems ought to be captured by a valid biodiversity measurement. For example, the Group on Earth Observations Biodiversity Observation Network (GEOBON) introduced the "Essential Biodiversity Variables" classification (EBVs), which summarizes all the relevant aspects that need to be factored in when measuring biodiversity (Pereira et al. 2013). The EBV classification includes genetic diversity, phenotypic distinctiveness, physiology, distribution and many more, and it suggests that R & A ought to be understood as one of the *components*, or parts of, biodiversity.

⁴ If an evenly distributed community of one million species were to be annihilated except for 100 species, the D_{Gini} will drop from 0.999999 to 0.99 (1% drop)—not representing the real damage sustained by the community (Daly 2018, p.11). For a fuller philosophical review, see Crupi (2019) and Justus (2011).

⁵ This statement does not distinguish between various purposes of biodiversity measure, which might validate any of the indices above for a *specific* purpose. Assessing whether single indices can be validated for specific purposes is beyond the scope of this paper.

Understanding species richness and abundance *as parts of* biodiversity implies that if species richness and abundance indices only measure a part of biodiversity, it is not necessarily the case that a direct relationship holds between the properties of a component and the properties of the whole. Said differently, measurements of species richness and abundance might not be representative of the true biodiversity status because the extrapolation of the properties of a part to the whole is not guaranteed.

The mereological worry presented above would not emerge if, on the other hand, species diversity were understood as a *sign of*, or proxy for, biodiversity. A proxy is a measurand that can be quantified when the actual quantity of interest cannot be directly assessed. Then, inferences from the proxy to the actual object of measurement are justified because some causal mechanism is known to exist between the proxy and the quantity of interest. In other words, by measuring an adequate proxy, inferences can be made about the *measurand* of interest because the two values co-vary. Sarkar (2005) has called a measurable property of an ecosystem that stands for a biodiversity measurement a “surrogate”. Were species R & A adequate proxies for biodiversity, then R & A measurements would co-vary with the status of biodiversity, the actual measurand. However, this hypothesis has been empirically falsified: species richness does not co-vary with other factors critical to biodiversity that should be assessed in a valid measurement, such as phenotypic distinctiveness, abundance or disparity.⁶ Of course, richness and abundance measurements do not need to give all the relevant information about biodiversity to be useful to specific purposes, but ecologists want to develop additional methods to capture biodiversity as target phenomenon. It seems that measuring biodiversity exclusively as species count and abundance cannot play this role.

2.4 A Processual Approach to Measuring Biodiversity

Even if understanding *species* diversity in terms of counts or evenness has been a core value of biodiversity research since its heyday (Wilson 1988), ecologists and conservation biologists agree that a valid measure of biodiversity should include more than species and specimens inventorying, for both practical and conceptual reasons. As already mentioned above, for example, not all species have the same pivotal ecological function, and some taxa “score” higher in their contribution to diversity than others, by being unequal in terms of evolutionary history, ecological functions, aesthetic value. Accordingly, measurements that do not black box this information must be developed.

⁶ For example, Pierre Taberlet and colleagues (2012), using large-scale multispecies vegetation data, demonstrated that species richness does not co-vary with genetic diversity, another important aspect of biodiversity that should be accounted for in a biodiversity operationalization. Similarly, ecologist Helmut Hillebrand and colleagues argue that measurements of species richness are not informative enough of biodiversity fluctuations and crisis as they “changes in species richness and its underlying components do not necessarily correlate with species compositional turnover” (Hillebrand et al. 2018, p.171), which constitute a critical metric in ecological studies.

Emphasizing the value attributed to evolutionary history, alternative approaches to measuring biodiversity have recently been suggested. An increasingly-adopted method is grounded in phylogenetic systematics, a relatively new branch of cladistics which relies on quantifying evolutionary trajectories. Philosophers Christopher Lean and James Maclaurin, (2016) have been advocating for adopting an index that captures this type of information as the best operationalization of biodiversity. The idea is simple: the relational history that separates two species or populations represented in a cladogram (an diagram showing ancestral relationships) can be quantified in at least two ways, either i) by counting the past ancestral speciation events, or ii) by applying equations to calculate the length of the cladogram's branches (Lean & Maclaurin 2016, p.27).

The most commonly cited evolutionary history-based formula is Faith's Phylogenetic Diversity (PD) index (Faith 1992), which relies on method (ii). Faith's PD index is

$$PD_{Faith} = Bx \frac{\sum_i^B LiAi}{\sum_i^B Ai},$$

where B is the number of branches in a cladogram, L is the weighted mean of the branches' length and A is the average abundance of species included in that branch. PD_{Faith} applies to taxonomic resolutions below the species level (eg. populations) and dictates new directions as to how biodiversity should be assessed. Phylogenetic indices make instrumental use of data about species, but do not focus on taxa count, and rather conceive of biodiversity as the evolutionary *process* that leads to differentiation, operationalizing it as length of cladograms.⁷

Philosophers and conservationists have stressed how radical conceiving of biodiversity as a *process* is, instead of conceiving biodiversity as the *result* of said process (Maclaurin & Sterelny 2008). Phylogeny-based measurements are taken to represent the *evolutionary potential* in past and in future biological systems, and in this sense they depart from measurements based on species richness or specimen count, which instead bear on the actual *results* of the evolutionary process. Phylogenetic measures are not co-variant with species richness distribution measures (Pio et al. 2011), so recently more conservationists have endorsed conceptualizing of biodiversity as representing phylogenetic history and, accordingly, they have endorsed measuring evolutionary potential, rather than the more traditional species count (see also Hartmann & André 2013; Milot et al. 2020).

Interestingly, this shift from measuring extant richness to measuring populations' phylogenetic history or other processual aspects corresponds to what has historically been a paradigm shift in conservation. Protecting and preserving biodiversity is increasingly understood as a matter of understanding and protecting ecological processes and diversification patterns instead of focusing on the safety of a few

⁷ Faith's PD index is only one of the many measures of biodiversity conceived as phylogenetic distance. These measurements are made possible because of the increasing availability of molecular and genetic data (see Vellend 2011)

charismatic species (Takacs 1996, p. 67-ff; Barnosky et al. 2017; Odenbaugh 2021, Soulé 1985).

Biodiversity researchers are unsatisfied by operationalizations of biodiversity centered around species and seem to be moving towards measurements that represent evolutionary potentialities.

3 Measuring paleodiversity

Developments in contemporary approaches to measuring biodiversity reflect the difficulty of conceptualizing biodiversity without reducing or reifying it to species richness or abundance. I concluded Section 2.4 by saying that more adequate operationalizations of biodiversity today should include metrics that quantify, for example, the evolutionary potential of lineages. Measuring biodiversity today is challenging, but measuring past biodiversity is also difficult. Nonetheless, a common mantra in biodiversity research is that data about past diversity are critical to quantify the exact status of biodiversity today, and whether we are currently in a crisis. Ecologist Helen Morlon and colleagues, for instance, write “Inferring rates of speciation and extinction and the resulting pattern of diversity over geological time scales is one of the most fundamental but challenging questions in biodiversity studies” (Morlon et al. 2011, p. 16327). In a similar vein, paleontologist Anthony Barnosky contends that a lot of what we should think about the current biodiversity crisis hinges on our ability to conduct “meaningful comparisons between modern conditions and long-term histories” (Barnosky et al. 2017, p.2). In this section, I focus on estimates of paleodiversity and past diversity fluctuations, to show the departure from contemporary methods of measuring what is supposed to be the same measurand, namely the status of biodiversity in the deep past.

Despite a large body of philosophical work devoted to biodiversity, its “deep time” dimension, namely paleodiversity, is a relatively under-explored topic in philosophy.⁸ This is unfortunate, since a deeper investigation into the conceptual, methodological, and epistemic assumptions in the study of past diversity could potentially contribute to the adage “the past is a guide to the future”.

Paleodiversity (a contraction of paleobiodiversity) is usually conceptualized as a representation of macroevolutionary patterns of diversity over time. These fluctuations track extinctions and speciation events that can, but do not necessarily, correspond to ecological and climatic disruption (Racki 2021). Having correct estimates of paleodiversity, and knowing that sometimes diversity fluctuations corresponded to climatic conditions, might be used to obtain more accurate predictions about the dynamics of current systems. Using the distinction introduced in Section 2.4, many paleodiversity studies conceptualize and measure past diversity not as a process, but as the result of evolutionary mechanism in terms of speciations and extinctions, meaning how many species or higher taxa obtain at various points in the past.

⁸ Some exceptions are Bokulich (2021) and Watkins (2021)

If a representation of macroevolutionary processes is in the scope of paleodiversity measurement, it is no surprise that these estimates necessarily rely on the fossil record (Benson and Mannion 2012). The information about the past is inferred from what remains of proxies for ancient forms of life, ecosystems and climatic conditions. Paleodiversity is commonly represented as a paleodiversity curve, such as the famous Sepkoski's diversity curve (Raup & Sepkoski 1982) or the more recent Phanerozoic diversity curve (Alroy et al. 2008) updated and corrected with metadata stored on the Paleobiology Database (Fig. 1).

Measuring paleodiversity happens in two stages. The first stage consists in, first and foremost, collecting fossil data in the field or from physical archives or, alternatively, retrieving data from various online databases, such as the Paleobiology Database. At this stage, a paleontologist has access to biased and fragmentary data about taxa and abundance across time. Just from this data, a "raw taxic diversity" estimate could be calculated, yet it would be extremely inaccurate. As philosopher Alisa Bokulich (2021) has pointed out, analytical approaches are routinely applied to adjust for various sources of the bias in fossil data. Therefore, the second stage of measuring paleodiversity consists in correcting data using various statistical techniques that will ultimately result in a more accurate paleocurve. This paleocurve displays taxonomic diversification and extinction patterns in the deep past that "pure" fossil data cannot show.⁹

Paleodiversity curves have the advantage of diachronically representing the trends and patterns in taxa fluctuations. Models have been developed to introduce some measure of taxa evenness in the estimate (Alroy et al. 2008). Paleocurves have been used, among others, to infer extinction and diversification events, and responses of fauna and flora to environmental disruptions. For instance, based on the fossil record and using sophisticated modeling techniques, five major extinction events and adaptive radiations have been described (Raup & Sepkoski 1982), a count that has long been contested (Racki 2021). Once an accurate paleodiversity curve is obtained, some argue, comparing patterns of paleodiversity and biodiversity are possible, and should inform us about the severity of the ongoing environmental emergency. In Part Section 4, I will argue that this inference from paleodiversity to a contemporary unspecified biodiversity crisis is unjustified.

More controversial, but potentially critical to justify the more specific inference that we are facing a taxonomic crisis, is to use paleodata to measure the background extinction rate of taxa in the past, normally presented as the number of species extinctions per million species per year (E/MSY, see Ceballos et al. 2015). Rates of extinction for today's threatened species have been suggested by E.O Wilson (1988), the International Forum on the Conservation of Nature (IUCN), and others. An accurate measurement of past extinction and speciation rate, corrected for sampling

⁹ Bokulich also argues that paleocurves are better thought of not as representations of taxa fluctuations obtained from "pure data". Rather, they should be conceived of as "data models", namely as the outcome of filtering and improving inevitably incomplete data using post-collection techniques that have been shown to be effective. This improved data displays higher degree of fidelity to the phenomenon of interest and is more useful to paleontologists than raw taxic diversity estimates.

biases and time-spans considerations, would be an adequate way of comparing the past rates of extinction to the present with possible implications for conservation plans. What is first needed for an accurate past extinction rate, though, is a reliable way of detecting, classifying, and counting extinct species. The same criteria hold for obtaining an accurate paleoecore. I now turn to the practical and then the conceptual issues involved in this operation.

3.1 Issues in paleodiversity data and measurements

Paleontological estimates are not only informed but also practically *constrained* by the fossil record, which is usually said to be incomplete and biased (Woolley et al. 2022).¹⁰ For instance, only a small fraction of the living world fossilizes in a process called “differential preservation”, with preference to those organisms with an exoskeleton or shell. In addition, few environmental conditions grant fossilization, like anoxic and fast-deposited locations: as a result, our paleocurves mostly reflect abundant organisms living in shallow marine environments. Additionally, the effort to collect fossil data is normally not chronologically or spatially homogeneous, which may result in sampling bias.

In order to polish fossil data and make them usable for specific purposes (e.g. to verify taxa-specific conjectures or local ecological hypotheses), successful statistical techniques have been developed since at least since the 1980s. Nonetheless, most paleontologists are still suspicious about extrapolating from paleocurves to global taxa fluctuations or about the efficacy of using paleocurves to say something about the overall status of taxonomic diversity today (Sepkoski 2020, ch. 4-5). Perhaps surprisingly, these type of considerations put pressure on the popular narrative according to which the Earth has experienced a fixed number of mass extinctions and is possibly entering a new catastrophic extinction event, a still ongoing debate among paleoecologists (see, for instance, Racki 2021; Plotnick et al. 2016; Barnosky et al. 2011).¹¹

¹⁰ The old trope in paleontology to assert that the fossil record is not an exhaustive or reliable source of evidence is still alive. This is unfortunate because eminent studies oppose this view and demonstrate that the fossil record can adequately serve specific purposes. For example, Gould and Eldredge most famously defended the completeness of the fossil record and used it to overturn traditional views on the mode and tempo of macroevolutionary patterns (Gould and Eldredge 1972). Furthermore, sophisticated analytical techniques have been developed to deal with incompleteness (Kidwell and Holland 2002), such as reconstructing phylogenetic trees for specific lineages (Woolley et al. 2022).

¹¹ It should be made clear here that talking about an ongoing biodiversity crisis and talking about entering a mass extinction are two separate issues. A mass extinction event is usually characterized by high rates of species extinction unbalanced by species origination or by a significant increase over the background extinction rate. Bocchi et al. (in press) provides a meticulous philosophical analysis of the concept of mass extinction and the evidence needed to support the claim that we are entering an extinction event comparable to those of the past). A biodiversity crisis is a more complex and multifaceted issue to assess than a mass extinction event. As I have already hinted at and as I will make explicit in Section 4, the former requires a type of evidence that is not reducible to evidence for an escalation of species extinction.

The adequacy of paleocurves to extrapolate general information (crosscutting taxa and at a global scale) has been subject of debate due to other practical and conceptual considerations that might additionally undermine the accuracy of taxonomic fluctuations. First, there are several practical problems in calculating the macroevolutionary patterns with an accurate degree of certainty. Some of the issues for drawing accurate fluctuation graphs comes from the *data journey*, *sensu* Leonelli (2016), that fossil data undergo from the collection site, to their cataloguing, to their digitalization. The datasets, compedia and lists providing taxonomic information for the fluctuation models are slowly being converted into web archives. Databases such as Sepkoski 1992, the one that first materialized the idea of biodiversity in deep time, are being digitalized, but the process might take decades, slowing down what can be inferred about paleodiversity. What is more, databases contain errors. One of the omnipresent errors in physical and digital catalogues, for example, is synonymity, i.e. when a single taxon is listed under at least two names. When data are corrected and taxa are reduced by removing the synonyms, the estimates about fluctuations and extinction rates might differ significantly.

Another issue for accurately measuring paleodiversity is fossil classification, which usually follows the criteria of the morphospecies concept, according to which organisms are clustered together based on shared phenotypic traits. The morphospecies concept alone is notoriously unable to distinguish cases of divergent or convergent evolution, and cladograms are highly underdetermined if based purely on that criterion. Sometimes, specimens that are morphologically similar are split in two or more species because they are found in separate stratigraphic layers (Gingerich 1979). Paleotaxonomists have taken stratigraphic interruptions as evidence that morphologically similar specimens cannot belong to the same species. This assumption is controversial and poses problems when calculating the background extinction rate.

New information regarding developmental considerations has also modified paleontologists' analysis of past animal taxonomies. Fossilized pieces have long been treated as belonging to adults following the assumption that immature bones are less likely to fossilize, so that skeletons and fragments actually belonging to the same species at two different developmental stages, have been classified as two distinct species (de Ricqlés et al. 2008). A more exhaustive analysis of the sources of uncertainty in measuring paleodiversity is beyond the scope of this paper.

As we will see, the types and nature of data involved in paleodiversity research play a key role in the incommensurability problem. These data add up to other conceptual issues that make claim such as “we are living in a general biodiversity crisis” ultimately poorly justifiable.

4 The Incommensurability Problem

In Sections 2 and 3, I analyzed how contemporary biodiversity and paleodiversity are measured, focusing on how the different measurements stem from their respective conceptual frameworks and require specific data for a proper assessment. In this part, I elaborate on what I call “the incommensurability problem” emerging from

attempts at comparing biodiversity measures and paleodiversity estimates to assess the severity of the ongoing “biodiversity crisis”. This incommensurability explains why loose claims that we are living in a biodiversity crisis lack strong justificatory support if the evidence required is related to paleodiversity.

As Paul Edwards has argued in the case of the climate crisis:

“to say that the global climate has changed implies that we know what it used to be. At a minimum, we are comparing the present with some period in the past. We would like to know the details, the trend over time [...] ideally 100 years or more. And since we are talking about global climate, we need some kind of picture of the whole planet [...]” (Edwards 2010, p.4)

In this section, I argue that, similarly to Edward’s example, to justify that biodiversity is currently undergoing a crisis, one must compare it to some known past state of non-crisis. But what usually is said to play the role of second term in the comparison, namely, paleodiversity, is conceptualized and quantified in a way that is incommensurable with how biodiversity is conceptualized and quantified presently, putting pressure on the claim that we have sufficient evidence to conclude that we are experiencing a biodiversity crisis. Similarly to the climate case, moreover, comparative claims between the present status of biodiversity and paleodiversity need to be based on relatively-accurate estimates that account for the same *type of data* (taxa, taxonomic resolution, evolutionary processes, temporal and spatial scales). This is also not the case.

Let me first pause to address a possible skeptical remark before developing my argument for incommensurability. One might argue that the inferential strategy to move from paleocurves to justify that we are living in a biodiversity crisis is a historic fiction and that it has never been seriously endorsed in scientific circles except for the few instances I mentioned previously. But in his recent book *Catastrophic Thinking* (2020), David Sepkoski has cogently shown that this tactic has a long history in scientific circles. According to Sepkoski, since the 1980s, biologists interested in conservation have become aware that it was “rhetorically effective to compare the current depletion of diversity to past mass extinctions, and even to predict that the anthropogenic species loss would eventually rival or exceed the greatest dying in the past” (Sepkoski 2020, p.234). Sepkoski refers to the justificatory use of past taxa fluctuation as “thoroughly entrenched” (Sepkoski 2020 p. 264) in the biodiversity crisis discourse. Early conservationists “calculated [the magnitude of the biodiversity crisis] by estimating the number of species extinctions in a given period (a day, a year, etc.) in relation to the number of species in existence and the magnitude of the problem [was] calculated by comparing current rates of extinction to those in the geologic past” (Sepkoski 2020, p.250). Sepkoski also reveals how this justification was not backed up by serious empirical evidence.¹² On their behalf,

¹² In Sepkoski’s historical reconstruction, E.O Wilson championed the past-present analogy to enhance his scientific agenda. Wilson and other environmentalists (Barbara Ward, René Dubos, Paul and Anne Erlich, David Ehrenfeld, Norman Myers) strongly advocated for the analogy regardless of their being aware of the speculative character of their estimates, as well as the protestations of paleontologists.

paleontologists would expect that a biodiversity crisis would comprise extinction patterns similar to past fluctuations, but they rejected this conclusion since it was based on too much uncertainty in present and past taxa estimates. Sepkoski does an excellent job in hinting at the disagreement in the scientific community about the kind of empirical evidence required to justify the claim that we are in a biodiversity crisis. However, he does not attribute this disagreement to the incommensurable conceptual roots of biodiversity and paleodiversity. I will argue below that ecologists and conservation biologists were overall at fault for not seeing the conceptual gap between biodiversity, which is irreducible to species count, and paleodiversity, which represents species fluctuation instead. Taken together, Sepkoski's argument and mine expose new facets of an old, influential, but fallacious argumentative strategy.

I here review some of the characteristics of biodiversity and paleodiversity that make them incommensurable, like when comparing apples with oranges instead of apples with apples. I discuss two sources of incommensurability: conceptual and data mismatch.

I take the conceptual incommensurability to be the most evident from the characterization of biodiversity and paleodiversity given above. By conceptual incommensurability I mean a lack of a common measure between the two concepts due to conceptual factors, such as definitional choices or ontological commitments. As I have shown in Sections 2.4 and 3, the criteria applied to measuring contemporary biodiversity are not the same criteria applied when estimating paleodiversity, due to different conceptual frameworks that guide what counts as a measurand, as well as how "biodiversity" and "paleodiversity" are operationalized according to these conceptual frameworks. To restate my point: paleodiversity measurements are meant to show macroevolutionary patterns, i.e., fluctuations in taxonomic richness. However, species inventories and diversity fluctuation do not exhaust the meaning of biodiversity as it is now conceptualized and measured. Biodiversity measures, especially those that calculate diversity as length of cladogram branches, result from conceptualizing biodiversity as the evolutionary process of biological systems.

The conceptual incommensurability stated here relates to using taxa count measurements in general. Even when estimating present extinction rates for the purpose of assessing the loss of biodiversity, there is a still a major conceptual problem: Is the species extinction rate a good indication of biodiversity loss? Is the reduction of biodiversity status to species number justified? For past biodiversity crises, biodiversity loss *just is* species loss; they are measured the same way in the fossil record: as rise and fall of taxa numbers. This is not the case for contemporary biodiversity, in which species are weighted for their functionality in specific ecosystems and their evolutionary history, and it might not be the case that a black-boxed species loss is either necessary or sufficient to signal biodiversity loss.¹³

¹³ Species loss is a definitional feature of a mass extinction, which are catastrophic events measured on a global scale during which an meaningful portion of extant species went extinct. For a biodiversity crisis to be occurring, this condition does not necessarily need to be met (Racki 2021).

Paleodiversity estimates are therefore underpinned by taxonomic concerns in a sense in which contemporary biodiversity estimates are not. To capture this tension, I will call those measures of diversity that are informed by a conceptual framework centered on taxa alone “taxa-based”: paleodiversity is in fact a measure of the *results* of evolutionary processes. In contrast, biodiversity measurements such as Faith’s PD index are underpinned by a “process-based” framework, which increasingly echoes the concern of quantifying *processes* themselves.

The gap in the theoretical frameworks that drives diversity research mirrors the different research programs responsible for making said measurements, as well as the purposes (*sensu* Bokulich and Parker 2021) that contemporary biodiversity measures and paleodiversity estimates adequately serve. Measuring biodiversity falls within the scope, in general, of conservation biology. Conservation biology is a future-oriented enterprise, insofar as conservationists are concerned about maintaining certain ecological functions more than about which species are performing said functions. Conservation, in its heyday, was an enterprise meant to target species, and it was motivated to protect all or almost all species from extinction. This focus was soon abandoned as impractical and conceptually flawed. Preserving all species is indeed extremely demanding in terms of research and resource allocation. Additionally, the idea of preserving species, knowing that species naturally go extinct and new species evolve, has been labelled “the paradox of conservation”. These considerations have all played a role in moving away from biodiversity conceived as taxa diversity, to biodiversity conceived as biological process. Species, under this framework, are understood as one of the many reifications of evolutionary processes. When measuring paleodiversity, on the contrary, the question is not what will happen but what already happened, namely which are the results of evolutionary trajectories and extinction/speciation patterns. For this reason, focusing on evolutionary processes rather than taxa count, is less intuitive in measuring paleodiversity. Not that paleodiversity research is not meant to guide conservation action or to make prediction about the magnitude of extinction events—this is the scope of a relatively new discipline called conservation paleoecology. But, overall, paleodiversity *per se* is conceptualized as representing patterns of speciation and extinction, where species are the units of interest. Accordingly, the operationalizations of the concepts of biodiversity and paleodiversity hinge on the purpose that the measurements serve. If the purposes differ, it comes as no surprise that the metrics themselves are incommensurable.

Let me be more specific about the interplay of biodiversity and paleodiversity measurements. I am not suggesting that the integration of contemporary biodiversity data obtained from a processual framework, such as measurements of the evolvability of traits, cannot and will not be implemented in paleodiversity estimates in general. Nor am I suggesting that there is no role for paleodiversity research in biodiversity research and conservation. Just the opposite is true: in the last few years, increasing attention has been given to transforming paleodata into information about processes, such as the evolutionary potential of bio-systems (see Louys 2012). Nonetheless, the assumption that taxa are the favored units of paleodiversity remains, and it is at odds with contemporary biodiversity research (see Reydon 2019). All I am suggesting is that the way contemporary biodiversity is conceptualized and

measured, and at the way paleodiversity is now conceptualized and now represented, results in measurements that are incommensurable. Therefore, the relevant comparative judgements resulting from weighing biodiversity against paleodiversity measurements are epistemically problematic. Research programs and environmental policies that rely on such comparative inferences from past diversity to present and future diversity need to develop a stronger epistemic justification that copes with the incommensurability in the conceptual frameworks. I will attempt a solution in the next section, after discussing the second type of incommensurability.

The second source of incommensurability emerges if we consider the type of data involved in biodiversity or paleodiversity measurements. I will argue that insofar as biodiversity and paleodiversity data are not directly comparable without restricting the comparison to specific taxa or without specifying the ecological hypotheses under consideration locally, a second type of incommensurability, “data incommensurability”, emerges between paleodiversity and biodiversity measurements. Now, I will tackle some of the inconsistencies resulting from paleodiversity and biodiversity data availability.

Taxa representation and distribution are one of the main reasons why biodiversity data and paleodiversity data are not commensurate without qualifications. Due to differential preservation as well as heterogeneous sampling efforts, the large majority of data used to draw paleodiversity curves, both in synoptic studies like Raup & Sepkoski (1982) and Alroy et al. (2008) and in more fined-grained ones like Fan et al. (2020), comes from abundant fossilized species representative of shallow marine environment, with specific phenotypic traits like a shell or exoskeleton. Paleocurves more directly represent fluctuations in the richness and abundance of marine taxa, and they should not be erroneously treated as proxies to represent global paleodiversity swings—as sometimes happens. The extrapolation from paleocurves to any comprehensive and global status of paleodiversity is not so straightforward and generates a practical problem for comparison. Paradoxically, the status of biodiversity suffers a data asymmetry in the other direction. The International Union for Conservation of Nature (IUCN), which managed the largest biodiversity database for threatened species, has assessed around 100,000 species for extinction risk, but it is highly biased toward terrestrial species, whereas most databases of marine species are still “data deficient”. The estimate that 28% of global species is threatened with extinction is a projection from terrestrial ecosystems. Additionally, “how many species there are on Earth and in the oceans” (Mora et al., 2011) is still speculative, leading to measurement uncertainties about species inventories, phylogenetic relationships, and evolutionary potential.

Another evident reason why biodiversity and paleodiversity data are incommensurable is that they operate at different taxonomic levels. As we saw above, classification poses constraints on accurately calculating paleocurves. To overcome the uncertainty in fossil taxonomy, Forey et al., (2004) recommend that fossils used in generating paleodiversity curves should not represent species’ extinctions and

diversification, but should use classifications at the genus or even family level.¹⁴ Sepkoski's paleocurves (1982) and Alroy's Phanerozoic diversity curve (2008) do not represent species' macroevolutionary patterns, but richness variations of families and higher taxa. As Abigail Lane and Michael Benton (2003) have noticed, it is unclear how the taxonomic level at which the paleocurves are built determines the pattern exhibited by the curve.¹⁵ It does not follow necessarily that the decline or increase in families and genera co-varies with decline or increase in species count. This is of course problematic to those who argue that evidence for the contemporary biodiversity crisis can be found in species extinctions: we don't know species' extinction patterns from the paleocurves. Suppose we then reconsider extinction in terms of higher taxa, like genera: Since 1500, around 800 extinction events at the species level have been documented (Ceballos et al. 2015) but only a handful of genera extinctions.¹⁶ The IUCN, for example, which handles the most authoritative database on extinction risk, does not explicitly report extinction at the level of genera or families.

Recently, paleontologists (such as David Raup in Sepkoski 2020 and Barnosky et al. 2011) and philosophers (Bocchi et al. *in press*) have focused on listing the various mismatches between paleodata and biodiversity data. For example, Bocchi and colleagues argue that taxonomic and geographic representativeness, theoretical commitments about which species concept to adopt, the choices during fossil preparation, and temporal resolution make inferences from paleodiversity estimates to the current status of biodiversity especially complex. These studies agree that mitigation strategies are needed if the two datasets are to be compared properly, and they advocate for the necessity of narrowing the scope of inferences that can be made from the past to the future to confirm more localized hypotheses.

A clarification is in order about my use of "incommensurability". "Incommensurability" usually alludes to Thomas Kuhn and Paul Feyerabend's idea that competing scientific theories lack common questions and methods, shared concepts and a general "world-view". Sometimes "paradigms" are said to be incommensurable, such as Ptolemaic and Newtonian physics, or Aristotelian and Darwinian biology.¹⁷ This

¹⁴ Some studies can quite accurately represent species fluctuations, but are generally rare and focused only on restricted numbers of taxa at lower spatial scales. See Fan et al., (2020)

¹⁵ Even if excellent work has been done in modelling how higher taxa track the diversity trends of lower taxa since the paleobiology revolution in the 1980s Grantham (2009), the strength of some of these methods in describing species patterns from genus or family patterns are successful only when an accurate identification of species, genus and family has been carried out (de Oliveira Jr et al. 2020, p.8). This excludes the possibility of applying some of these method to the fossil record, where higher taxa and not species are normally identified

¹⁶ I am grateful to Thomas Hegna for a useful email exchange about genera extinction.

¹⁷ A great deal has been written about what Kuhn and Feyerabend meant by "incommensurability", as well as about Kuhn's reformulation of the concept from *The Structure of Scientific Revolutions* (1970) to *The Road since Structure* (2000). A conceptual distinction, introduced by Sankey and Hoyningen-Huene (2001) is that between semantic (or taxonomic) and methodological incommensurability. Two theories are said "semantically incommensurable" when the theoretical terms they employ cannot be co-opted in other theories, because their definition is embedded within the specific context within which that concept was developed (see Sankey and Hoyningen-Huene 2001, p. ix). The semantic aspect of incommensurability is the most commonly mentioned in philosophy. Two theories are said "methodologically incom-

paper is not about theories but about concepts and their operationalization (even if concepts and measurements normally operate within theories). Accordingly, the meaning of “incommensurability” adopted here might differ from traditional use. By “incommensurability” I simply mean “lack of a common measure” between the concepts of biodiversity and paleodiversity, and, necessarily, between the measurements adopted to quantify them.¹⁸ The incommensurability between paleodiversity and biodiversity will describe the absence of shared conceptual criteria and relevant metrics that make their values impossible to compare (Box 1).

BOX 1

Incommensurability between Biodiversity and Paleodiversity:

Conceptual Incommensurability: the way biodiversity is conceptualized and the way paleodiversity is conceptualized lack a common measure.

Data Incommensurability: the data collected to accurately inform paleoestimates are different types of data than are needed for biodiversity estimates.

To sum up this section: The form of my argument has been that, when we make comparative claims between the past and present, these claims must be based on shared criteria to quantify the two terms of the comparison. In the case of biodiversity and paleodiversity we lack those criteria for both conceptual and practical reasons. Therefore the comparison is unjustified. If the comparison is unjustified, unqualified claims such as “we are living in a biodiversity crisis”, usually supported by appealing to a comparison between biodiversity and paleodiversity, lack a strong justificatory basis.

5 Three possible solutions

So far, I have argued that a comparative judgement between present biodiversity and paleodiversity is epistemically hard to justify, given that the metrics used to assess the two compared values, as well as the conceptual framework that motivates said measurements, are incommensurable. If it all boils down to different conceptualizations of what is supposed to be the same measurand, namely current and past diversity, then there seem to be an obvious solution to the incommensurability problem. This solution consists of either redefining biodiversity and paleodiversity (or both)

Footnote 17 (continued)

measurable” if they abide by different standards when it comes to adjudicating which, of two competing theories, best fits the data (see Sankey and Hoyningen-Huene 2001, p. xiii).

¹⁸ This meaning of incommensurability is still bound to that Kuhn adopted in *Structure* as a holistic concept expressing a possible relationship between two heterogeneous domains (see also Oberheim & Hoyningen-Huene 2018; Patton 2015).

until they significantly converge, so that a unique measurement procedure could be validated and comparison between their values can be carried out. This can happen in two ways.

First, the concept of paleodiversity could be restructure so as to necessarily comprise phylogenetic distance or information about functionality, and it could be measured with something that is not the paleodiversity curve. This is an ambitious solution whose viability is an empirical matter. So far, this project is still in the making and it might be long before it is implemented in paleodiversity research.

Second, and alternatively, one might argue that the concept of biodiversity is overly demanding, and what we really mean by “biodiversity crisis” is actually taxa loss. Granting this assumption, a comparison of biodiversity to paleodiversity is potentially viable (insofar as there is a homogeneous taxonomic resolution). I suspect this solution would be unpopular: the reification of biodiversity as taxa count would violate the most recent trends in biodiversity research, which generally agree in rejecting the operationalization of the biodiversity concept to species count. Therefore this second solution seems to me to be less preferable than the first.

I will detail a third way of tackling the incommensurability problem inspired by Carlos Santana’s (2014) “biodiversity eliminativism”. This third alternative aims at eliminating talk of a general “biodiversity crisis” once and for all and instead breaks down the concept into its components (phylogenetic history, taxa richness and abundance, ecosystem services, etc.), each of which is more easily quantifiable and potentially comparable to paleodata.

The philosophical literature is not new to eliminativist arguments when it comes to the concept of biodiversity. Most notably, Santana (2014) adopted a form of biodiversity eliminativism to answer the question of what is the object of conservation. In his provocative paper “Save the Planet: Eliminate Biodiversity”, Santana suggested eliminating the word “biodiversity” from conservation agendas, and to focus on “biological values” instead. For him, “biodiversity” captures an increasingly expanding and complex area of research, as has been argued in Section 2 which fails to be “straightforwardly operationalizable” (Santana 2014, p.763). “Biological values” instead refers to anything conservationists want to preserve—species richness at a specific location, the functionality of an ecosystem, a unique historical trajectory embodied by a genus, etc. Biological values are instrumental to some purposes and can be operationalized more easily and flexibly than the concept of biodiversity.

Following Santana, the indices for biodiversity listed in Section 2 are not to be interpreted as measurements of biodiversity *per se*, but they can be co-opted to assess multiple biological values. This move looks *prima facie* similar to the attempt at measuring biodiversity by measuring one of its proxies that I mentioned in Section 2.3. But Santana’s account differs from it significantly, insofar as it eliminates any need for conceptualizing “biodiversity” and operationalizing it over and above some measurable and discrete environmental properties. This solution advocates for a form of eliminativism with respect to assessing biodiversity as a unique *measurand*.

Santana’s account could be adapted to solve the incommensurability problem as it enacts a restriction in the scope of inferences from the past to the present.

Using Santana's view, the justification of claims such as "we are in the middle of a biodiversity crisis" does not hinge on a proper comparison between paleodata *per se* and the current state of biodiversity *per se*. There is no need to talk about an overall "biodiversity crisis" at all. Rather, "crisis" amounts to a loss of a biological value, such as the disruption of a specific ecosystem services, or the predicted loss of a specific hotspot. These dimensions are more easily measured and do not hinge on any complex, multilayered idea of biodiversity. Additionally, the threshold for concern that stands for a "crisis" in any of these biological values is recognized as intrinsically anthropogenic and semi-arbitrary.

An example might help. If the biological value of interest to a study is the species richness at a specific location, and the conservation effort aims at preserving such richness, the data about past and present trends in species richness—even if not representative of biodiversity in its entirety—will be epistemically sufficient to make comparative judgements. There is no need to appeal to any abstract, all encompassing idea of biodiversity, if the biological value is species richness. All you need is the right type of data. Finally, comparative claims such that species richness in that specific location is declining can potentially be justified using evidence from the past, or just by making reference to anthropogenic criteria. This does not amount to support for the stronger claim that we are facing a global biodiversity crisis, a claim that is not so central anymore, once Santana's suggestion has been adopted

The eliminativist solution presented here has three main benefits:

1. Potential efficacy in informing conservation strategies without necessarily relying on vague or inaccurate biodiversity and paleodiversity estimates, the result of operationalizing two broad and incommensurable concepts.
2. It narrows the scope of comparative claims requiring specific types of evidence that are potentially extractable from the fossil record or other environmental proxies.
3. It allows for semi-arbitrary threshold to demarcate a state of crisis without necessarily relying on deep-time standards.

This eliminativist solution might not be better *in principle* than the first one mentioned above, but it seems to me better *pragmatically* and readily applicable to conservation work. More accurate policies can be developed and justified once a deconstruction of biodiversity is operationalized and its status is assessed using measurements that can, but do not need to, be compared to paleodata.

Accepting a form of eliminativism to solve the conceptual incommensurability between biodiversity and paleodiversity measurements might come at the cost of losing some sympathizers. This third strategy could be accused of problematically eliminating any need to justify the global-scale hypothesis that we are experiencing an overarching "biodiversity crisis", which may in fact be dangerously occurring. I might be accused of preferring the security of strong comparative inductions that demonstrate some biological values are deteriorating to the precautionary principle, according to which it is better to be safe and admit that we

are experiencing a biodiversity crisis, than to be sorry. But if the incommensurability problem is correct, this criticism seems to be at odds with the premises that biodiversity is currently conceptualized and measured in a complex way and that a state of crisis should be justified with adequate empirical evidence. I see it as problematic to claim that a decline, possibly crosscutting all levels of biodiversity analysis, is happening without specifying the kind of evidence that is needed for this claim as well as the evidence we possess. It might be the case that the measurements of all biological values we can think of will eventually prove to be declining, therefore confirming what I think this criticism seem to care about. But we ought not wait, either to affirm with certainty or to take action, until we have all this evidence to be sure that some biological values are declining.

I can also see opponents to this third solution objecting that eliminativism may hurt the rhetorically-effective discourse around an anthropogenic biodiversity crisis happening on a global scale. This rhetoric-based criticism might be grounded on the fear that eliminativism is nothing but a form of skepticism toward the ability of scientists to provide pertinent evidence to support general ecological hypotheses. This skepticism might be interpreted by some as a first step toward denialism about the danger of losing biodiversity. I think these worries are illegitimate and disregard a positive implication of adopting an eliminativist strategy, namely that science can do better than stopping at a bird's-eye view. Biodiversity eliminativism focuses on the ability of biodiversity research to test more accurate hypotheses and possibly make more accurate predictions once the pernicious rhetoric of an unqualified biodiversity crisis is abandoned. More accurate claims can support better informed conservation strategies and open up new paths of more rigorous science.

It should be emphasized that the argument presented for eliminativism does not try to undervalue the danger of possible crises in biodiversity values. Nor do I want to make the case that the phrase should be banned for having mostly a rhetorical purpose. What interests me is the justification that seems to emerge from scientific and popular literature, namely that one can draw conclusions about the status of biodiversity today, whether we are in a crisis, by operating an unqualified comparison to paleodiversity fluctuations. I argue that this inference is doubtful and the problem of incommensurability should be disseminated loudly in philosophical and scientific circles alike.

6 Conclusion

I argued that paleodiversity and biodiversity measurements are underpinned by different theoretical frameworks. This results in biodiversity measurements that are incommensurable with paleodiversity measurements. Therefore, comparative judgments and inferences about the present status of biodiversity based on comparison with paleodiversity are weakly justified. I then zoomed in on what I take to be the most viable solution: Adopting Santana's suggestion of eliminating biodiversity and concentrating on measuring any environmental feature of value that conservation biology strives to preserve, without relating that measurement to a broader concept of biodiversity. This significantly reduces the scope of comparative claims, as well

as the epistemic justifications required. Following my suggestion, the claim that we are experiencing a species richness crisis, or a decline in a specific ecosystem function, can be strengthened by paleodata, but the same paleodata cannot significantly support the more abstract claim that we are in a global biodiversity crisis.

Acknowledgements I am grateful to Alisa Bokulich, Aja Watkins and the Phi-Geo research group, as well as the audience at the 2021 EPSA conference in Turin. I am especially grateful to Emily Parke, Lydia Patton, and two anonymous reviewers for useful feedback on a previous version of this manuscript.

Declarations

Ethical Statement The author declares no conflict of interest. This research did not receive any financial support. No human or animal trials were performed in this study.

References

- Alroy, J., Aberhan, M., Bottjer, D. J., Foote, M., Fürsich, F. T., Harries, P. J., Hendy, A. J. W., Holland, S. M., Ivany, L. C., Kiessling, W., Kosnik, M. A., Marshall, C. R., McGowan, A. J., Miller, A. I., Olszewski, T. D., Patzkowsky, M. E., Peters, S. E., Villier, L., Wagner, P. J., Bonuso, N., Borkow, P. S., Brenneis, B., Clapham, M. E., Fall, L. M., Ferguson, C. A., Hanson, V. L., Krug, A. Z., Layou, K. M., Leckey, E. H., Nürnberg, S., Powers, C. M., Sessa, J. A., Simpson, C., Tomašových, A., Visaggi, C. C. (2008). Phanerozoic Trends in the Global Diversity of Marine Invertebrates. *Science*, 321(5885), 97–100. Retrieved 2021-10-13, from <https://doi.org/10.1126/science.1156963>
- Barnosky, A. D., Hadly, E. A., Gonzalez, P., Head, J., Polly, P. D., Lawing, A. M., Eronen, J. T., Ackerly, D. D., Alex, K., Biber, E., Blois, J., Brashares, J., Ceballos, G., Davis, E., Dietl, G. P., Dirzo, R., Doremus, H., Fortelius, M., Greene, H. W., Hellmann, J., Hickler, T., Jackson, S. T., Kemp, M., Koch, P. L., Kremen, C., Lindsey, E. L., Looy, C., Marshall, C. R., Mendenhall, C., Mulch, A., Mychajliw, A. M., Nowak, C., Ramakrishnan, U., Schnitzler, J., Das Shrestha, K., Solari, K., Stegner, L., Stegner, M. A., Stenseth, N. C., Wake, M. H., Zhang, Z. (2017). Merging paleobiology with conservation biology to guide the future of terrestrial ecosystems. *Science*, 355(6325), eaah4787. Retrieved 2021-10-13, from <https://doi.org/10.1126/science.aah4787>
- Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., Marshall, C., McGuire, J. L., Lindsey, E. L., Maguire, K. C., Mersey, B., Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived?. *Nature*, 471(7336) 51–57, Retrieved 2021-10-13, from <https://doi.org/10.1038/nature09678>
- Benson, R. B. J., & Mannion, P. D. (2012). Multi-variate models are essential for understanding vertebrate diversification in deep time. *Biology Letters*, 8(1), 127–130. <https://doi.org/10.1098/rsbl.2011.0460>
- Bocchi, F., Bokulich, A., Castillo-Brache, L., Grand-Pierre & Watkins, A. (forthcoming). Are we in a sixth mass extinction? the challenges of answering and value of asking. *British Journal for Philosophy of Science*
- Bokulich, A. (2021). Using models to correct data: paleodiversity and the fossil record. *Synthese*, 198(S24) 5919–5940. Retrieved 2022-01-30 from <https://doi.org/10.1007/s11229-018-1820-x>
- Bokulich, A., & Parker, W. (2021). Data models, representation and adequacy-for-purpose. *European Journal for Philosophy of Science*, 11(1), 1–26.
- Ceballos, G., Ehrlich, P. R., Barnosky, A. D., García, A., Pringle, R. M., Palmer, T. M. (2015). Accelerated modern human-induced species losses: Entering the sixth mass extinction. *Science Advances*, 1(5) e1400253. Retrieved 2022-02-04, from <https://doi.org/10.1126/sciadv.1400253>
- Crupi, V. (2019). Measures of biological diversity: Overview and unified framework
- Daly, A., Baetens, J., De Baets, B., (2018) Ecological Diversity: Measuring the Unmeasurable. *Mathematics*, 6(7)119, Retrieved 2021-10-13, from <https://doi.org/10.3390/math6070119>

- de Oliveira Jr, S. S., Ortega, J. C. G., dos Santos Ribas, L. G., Lopes, V. G., & Bini, L. M. (2020). Higher taxa are sufficient to represent biodiversity patterns. *Ecological Indicators*, *111*, 105994.
- de Ricqlés, A., Padian, K., Knoll, F., Horner, J. R. (2008) On the origin of high growth rates in archosaurs and their ancient relatives: Complementary histological studies on Triassic archosauriforms and the problem of a “phylogenetic signal” in bone histology. *Annales de Paléontologie*, *94*(2)57–76. Retrieved 2022-02-04, from <https://doi.org/10.1016/j.annpal.2008.03.002>
- Edwards, P. N. (2010) A vast machine: computer models, climate data, and the politics of global warming. Cambridge, Mass: MIT Press. (OCLC: ocn430736496)
- Faith, D. P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*(1) 1-10. Retrieved 2021-10-13, from [https://doi.org/10.1016/0006-3207\(92\)91201-3](https://doi.org/10.1016/0006-3207(92)91201-3)
- Fan, J.-X., Shen, S.-Z., Erwin, D. H., Sadler, P. M., MacLeod, N., Cheng, Q.-M., Hou, X.-D., Yang, J., Wang, X.-D., Wang, Y., et al. (2020). A high-resolution summary of cambrian to early triassic marine invertebrate biodiversity. *Science*, *367*(6475), 272–277.
- Forey, P. L., Fortey, R. A., Kenrick, P., Smith, A. B. (2004). Taxonomy and fossils: a critical appraisal. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* *359* (1444) 639–653, <https://doi.org/10.1098/rstb.2003.1453>,
- Gingerich, P. D. (1979). The Stratophenetic Approach to Phylogeny Reconstruction in Vertebrate Paleontology. In J. Cracraft & N. Eldredge (Eds.), *Phylogenetic Analysis and Paleontology* (pp. 41-78). Columbia University Press. Retrieved 2022-02-04, from <https://doi.org/10.7312/crac92306-004>
- Gini, C. (1912). *Variabilità e mutabilità (variability and mutability)* (p. 156). Bologna: Cuppini.
- Gleason, H. A. (1922). On the relation between species and area. *Ecology*, *3*(2), 158–162.
- Gould, S. J., & Eldredge, N. (1972). Punctuated equilibria: an alternative to phyletic gradualism. *Models in paleobiology*, *1972*, 82–115.
- Grantham, T. (2009). Taxic paleobiology and the pursuit of a unified evolutionary theory. *The Paleobiology Revolution*, 215-238.
- Hartmann, K., & André, J., (2013). Should evolutionary history guide conservation? *Biodiversity and Conservation*, *22* (2), 449-458. Retrieved 2021-10-13, from <https://doi.org/10.1007/s10531-012-0422-z>
- Hillebrand, H., Blasius, B., Borer, E. T., Chase, J. M., Downing, J. A., Eriksson, B. K., Filstrup, C. T., Harpole, W. S., Hodapp, D., Larsen, S., Lewandowska, A. M., Seabloom, E. W., Van de Waal, D. B., Ryabov, A. B. (2018). Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. *Journal of Applied Ecology*, *55* (1), 169-184. Retrieved 2022-02-01, from <https://doi.org/10.1111/1365-2664.12959>
- Justus, J. (2011). A Case Study in Concept Determination: Ecological Diversity. In: *Handbook of The Philosophy of Science: Philosophy of Ecology* (pp. 147-168). San Diego: North Holland: Elsevier.
- Kidwell, S. M., & Holland, S. M. (2002). The quality of the fossil record: implications for evolutionary analyses. *Annual Review of Ecology and Systematics*, *33*(1), 561–588.
- Kuhn, T. S. (1970). The structure of scientific revolutions (Vol. 111). Chicago, University of Chicago Press.
- Lane, A., & Benton, M. (2003). Taxonomic level as a determinant of the shape of the phanerozoic marine biodiversity curve. *The American Naturalist*, *162* (3), 265-276. Retrieved 2021-10-13, from <https://doi.org/10.1086/377188>
- Lean, C., & Maclaurin, J. (2016). The value of phylogenetic diversity. In *Biodiversity conservation and phylogenetic systematics* (pp. 19-37). Springer, Cham.
- Leonelli, S. (2016). *Data-centric biology: a philosophical study*. Chicago ; London: The University of Chicago Press.
- Louys, J. (Ed.). (2012). *Paleontology in ecology and conservation*. Berlin Heidelberg: Springer.
- Maclaurin, J., & Sterelny, K. (2008). *What is biodiversity?* Chicago: University of Chicago Press. (OCLC: ocn156874812)
- Magurran, A. E., & McGill, B. J. (Eds.). (2011). *Biological diversity: frontiers in measurement and assessment*. Oxford ; New York: Oxford University Press. (OCLC: ocn642283168)
- Malaterre, C. (2013). Microbial diversity and the “lower-limit” problem of biodiversity. *Biology & Philosophy*, *28*(2), 219–239.
- Margalef, R. (1958). Information theory in biology. *General Systems Yearbook*, *3*, 36–71.

- Menhinick, E. F. (1964). A comparison of some species-individuals diversity indices applied to samples of field insects. *Ecology*, 45 (4), 859–861. Retrieved 2021-10-13, from <https://doi.org/10.2307/1934933>
- Milot, E., Béchet, A., & Maris, V. (2020). The dimensions of evolutionary potential in biological conservation. *Evolutionary Applications*, 13 (6), 1363–1379. Retrieved 2021-12-05, from <https://doi.org/10.1111/eva.12995>
- Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How Many Species Are There on Earth and in the Ocean? *PLoS Biology*, 9 (8), e1001127. Retrieved 2021-10-13, from <https://doi.org/10.1371/journal.pbio.1001127>
- Morlon, H., Parsons, T. L., & Plotkin, J. B. (2011). Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences*, 108 (39), 16327–16332. Retrieved 2022-02-03, from <https://doi.org/10.1073/pnas.1102543108>
- Oberheim, E., & Hoyningen-Huene, P. (2018). The incommensurability of scientific theories. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2018 ed.). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/fall2018/entries/incommensurability/>
- Odenbaugh, J. (2021). Conservation biology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2020 ed.). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2021/entries/conservation-biology/>
- O'Malley, M. (2014). *Philosophy of microbiology*. Cambridge University Press.
- Patton, L. (2015). Incommensurability and the bonfire of the meta-theories: Response to mizrahi. *Social Epistemology Review and Reply Collective*, 4(7), 51–58.
- Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H. M., Cardoso, A. C., Coops, N. C., Dulloo, E., Faith, D. P., Freyhof, J., Gregory, R. D., Heip, C., Hoft, R., Hurtt, G., Jetz, W., Karp, D. S., McGeoch, M. A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J. P. W., Stuart, S. N., Turak, E., Walpole, M., Wegmann, M. (2013) Essential biodiversity variables. *Science*, 339 (6117), 277–278. Retrieved 2021-10-13, from <https://doi.org/10.1126/science.1229931>
- Pio, D. V., Broennimann, O., Barraclough, T. G., Reeves, G., Rebelo, A. G., Thuiller, W., Guisan, A., & Salamin, N. (2011). Spatial predictions of phylogenetic diversity in conservation decision making. *Conservation Biology*, 25(6), 1229–1239.
- Plotnick, R. E., Smith, F. A., & Lyons, S. K. (2016). The fossil record of the sixth extinction. *Ecology Letters*, 19 (5), 546–553. Retrieved 2021-10-13, from <https://doi.org/10.1111/ele.12589>
- Racki, G. (2021). Big 5 Mass Extinctions. In *Encyclopedia of Geology* (pp. 603–616). Elsevier. Retrieved 2022-01-30, from <https://doi.org/10.1016/B978-0-12-409548-9.12028-7>
- Raup, D. M., & Sepkoski, J. J. (1982). Mass Extinctions in the Marine Fossil Record. *Science*, 215 (4539), 1501–1503. Retrieved 2021-10-13, from <https://doi.org/10.1126/science.215.4539.1501>
- Reydon, T. A. (2019). Are species good units for biodiversity studies and conservation efforts? In: *From assessing to conserving biodiversity* (pp. 167–193). Springer, Cham.
- Sankey, H., & Hoyningen-Huene, P. (2001). Introduction: Incommensurability and related matters. *Boston Studies in the Philosophy of Science*, 216, 7–34.
- Santana, C. (2014). Save the planet: eliminate biodiversity. *Biology & Philosophy*, 29 (6), 761–780. Retrieved 2021-10-13, from <https://doi.org/10.1007/s10539-014-9426-2>
- Santana, C. (2018). Biodiversity is a chimera, and chimeras aren't real. *Biology & Philosophy*, 33 (1–2), 15. Retrieved 2021-10-24, from <https://doi.org/10.1007/s10539-018-9626-2>
- Sarkar, S. (2005). *Biodiversity and Environmental Philosophy: An Introduction.*, 278
- Sepkoski, D. (2020). *Catastrophic thinking: extinction and the value of diversity*. Chicago: University of Chicago Press.
- Simpson, E. H. (1949). *Measurement of diversity*. *nature*, 163(4148), 688–688.
- Soulé, M. E. (1985). What is conservation biology? *BioScience*, 35(11), 727–734.
- Taberlet, P., Zimmermann, N. E., Englisch, T., Tribsch, A., Holderegger, R., Alvarez, N., Niklfeld, H., Coldea, G., Mirek, Z., Moilanen, A., Ahlmer, W., Marsan, P. A., Bona, E., Bovio, M., Choler, P., Cieślak, E., Colli, L., Cristea, V., Dalmás, J.P., Frajman, B., Garraud, L., Gaudeul, M., Gielly, L., Gutermann, W., Jogan, N., Kagalo, A. A., Korbecka, G., Küpfer, P., Lequette, B., Letz, D. R., Manel, S., Mansion, G., Marhold, K., Martini, F., Negrini, R., Niño, F., Paun, O., Pellecchia, M., Perico, G., Piękoś Mirkowa, H., Prosser, F., Puşcaş, M., Ronikier, M., Scheuerer, M., Schneeweiss, G. M., Schönswetter, P., Schrott Ehrendorfer, L., Schüpfer, F., Selvaggi, A., Steinmann, K., Thiel Egenter, C., Loo, M., Winkler, M., Wohlgemuth, T., Wraber, T., Gugerli, F. (2012). Genetic

- diversity in widespread species is not congruent with species richness in alpine plant communities. *Ecology Letters*, 15 (12), 1439–1448. Retrieved 2022-02-01, from <https://doi.org/10.1111/ele.12004>
- Takacs, D. (1996). The idea of biodiversity: philosophies of paradise
- Tal, E. (2013). Old and new problems in philosophy of measurement. *Philosophy Compass*, 8(12), 1159–1173.
- Thomas, K. (2000). The road since structure
- Vellend, M., Cornwell, W. K., Magnuson-Ford, K., & Mooers, A. (2011). Measuring phylogenetic biodiversity, 14
- Watkins, A. (2021). The Epistemic Value of the Living Fossils Concept. *Philosophy of Science*, 88 (5), 1221–1233. Retrieved 2022-01-30, from <https://doi.org/10.1086/714875>
- Wilson, E. O. (Ed.). (1988). *Biodiversity*. Washington, D.C: National Academy Press. (Meeting Name: National Forum on BioDiversity)
- Woolley, C. H., Thompson, J. R., Wu, Y. -H., Bottjer, D. J., Smith, N. D. (2022). A biased fossil record can preserve reliable phylogenetic signal. *Paleobiology*, 1–16

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.