

Bootstrapping outperforms community-weighted approaches for estimating the shapes of phenotypic distributions

Brian S. Maitner¹  | Aud H. Halbritter²  | Richard J. Telford²  | Tanya Strydom^{3,4}  |
 Julia Chacon⁵  | Christine Lamanna⁶  | Lindsey L. Sloat⁷  | Andrew J. Kerkhoff⁸  |
 Julie Messier⁹  | Nick Rasmussen¹⁰  | Francesco Pomati¹¹  | Ewa Merz¹¹  |
 Vigdis Vandvik²  | Brian J. Enquist⁵ 

¹Department of Geography, University at Buffalo, Buffalo, New York, USA; ²Department of Biological Sciences and Bjerknes Centre for Climate Research, University of Bergen, Bergen, Norway; ³Département de Sciences Biologiques, Université de Montréal, Montréal, Québec, Canada; ⁴Québec Centre for Biodiversity Sciences, Montréal, Québec, Canada; ⁵Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, USA; ⁶World Agroforestry Centre (ICRAF), Nairobi, Kenya; ⁷World Resources Institute, Washington, District of Columbia, USA; ⁸Department of Biology, Kenyon College, Gambier, Ohio, USA; ⁹Department of Biology, University of Waterloo, Waterloo, Ontario, Canada; ¹⁰California Department of Water Resources, West Sacramento, California, USA and ¹¹Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf, Switzerland

Correspondence

Brian S. Maitner
Email: bmaitner@gmail.com

Richard J. Telford
Email: Richard.Telford@uib.no

Funding information

Biological and Environmental Research, Grant/Award Number: DEAC02-05CH11231; National Science Foundation, Grant/Award Number: HDR-1934790 and BoCP-2225076; Norges Forskningsråd, Grant/Award Number: 287784; Watershed Function Scientific Focus Area and the ExaShed Project; Swiss National Science Foundation, Grant/Award Number: 182124

Handling Editor: Pablo Duchen

Abstract

1. Estimating phenotypic distributions of populations and communities is central to many questions in ecology and evolution. These distributions can be characterized by their moments (mean, variance, skewness and kurtosis) or diversity metrics (e.g. functional richness). Typically, such moments and metrics are calculated using community-weighted approaches (e.g. abundance-weighted mean). We propose an alternative bootstrapping approach that allows flexibility in trait sampling and explicit incorporation of intraspecific variation, and show that this approach significantly improves estimation while allowing us to quantify uncertainty.
2. We assess the performance of different approaches for estimating the moments of trait distributions across various sampling scenarios, taxa and datasets by comparing estimates derived from simulated samples with the true values calculated from full datasets. Simulations differ in sampling intensity (individuals per species), sampling biases (abundance, size), trait data source (local vs. global) and estimation method (two types of community-weighting, two types of bootstrapping).
3. We introduce the `TRAITSTRAP` R package, which contains a modular and extensible set of bootstrapping and weighted-averaging functions that use community composition and trait data to estimate the moments of community trait distributions with their uncertainty. Importantly, the first function in the workflow, `trait_fill`, allows the user to specify hierarchical structures (e.g. plot within site, experiment vs. control, species within genus) to assign trait values to each taxon in each community sample.

Brian S. Maitner, Aud H. Halbritter and Richard J. Telford contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

4. Across all taxa, simulations and metrics, bootstrapping approaches were more accurate and less biased than community-weighted approaches. With bootstrapping, a sample size of 9 or more measurements per species per trait generally included the true mean within the 95% CI. It reduced average percent errors by 26%–74% relative to community-weighting. Random sampling across all species outperformed both size- and abundance-biased sampling.
5. Our results suggest randomly sampling ~9 individuals per sampling unit and species, covering all species in the community and analysing the data using nonparametric bootstrapping generally enable reliable inference on trait distributions, including the central moments, of communities. By providing better estimates of community trait distributions, bootstrapping approaches can improve our ability to link traits to both the processes that generate them and their effects on ecosystems.

KEYWORDS

body size, community ecology, community-weighted mean, functional ecology, functional traits, nonparametric bootstrapping, population biology, R package, traitstrap

1 | INTRODUCTION

1.1 | Trait distributions

The shape and dynamics of phenotypic distributions can inform us of the rules of life that shape biological diversity and functioning across all scales of organization (Brown, 1995; Maurer, 1999). For example, at the population level, the shape of phenotypic distributions reflects how populations respond to different selective pressures and other evolutionary forces (Lande, 1976; Levins, 1968; Turelli & Barton, 1994). At the community level, phenotypic distributions can reflect species interactions, environmental pressures, immigration, local extinction and drive species coexistence and ecosystem functioning (Bolnick et al., 2011; Enquist et al., 2015; Keddy, 1992; Norberg et al., 2001). Their shapes reflect how assemblages have responded to past changes and how their composition and functioning will respond to future environmental changes (Grime, 2006; Inouye et al., 2019; Keddy, 1992; Suding et al., 2008; Webb et al., 2010). At larger geographic and temporal scales, phenotypic distributions reflect unique evolutionary histories, the impact of climate change, differing biogeographic processes and the unique climates and ecology of biomes (Donoghue & Edwards, 2014; Echeverría-Londoño et al., 2018; Moncrieff et al., 2016; Violle et al., 2014; Woodward et al., 2004). The 'true' trait distribution of a population or a species assemblage thus reflects the joint outcome of the ecological and evolutionary forces that have shaped the phenotypic diversity in that assemblage (Roughgarden, 1972, 1974, 1979; Figure 1, top row). However, measuring trait distributions is difficult, if not impossible, as it requires information on the trait values of all individuals (or organs, e.g. leaves) present. Biologists, therefore, rely on methods that estimate these distributions from the data we can feasibly sample.

The existing approaches for quantifying the trait diversity of an assemblage can be grouped into two categories: (1) those based on distance matrices and (2) those based on trait spaces (Mammola et al., 2021). Approaches based on distance matrices (e.g. functional dispersion; Laliberté & Legendre, 2010; functional distinctiveness; Violle et al., 2017) first convert trait observations into distance matrices (e.g. Euclidean for continuous traits, Gower for mixtures of continuous and categorical traits), and may further convert these distance matrices into functional dendrograms (e.g. Functional Diversity sensu Petchey & Gaston, 2002). Such distance-based approaches have the advantage of being comparable to other facets of biodiversity, such as phylogenetic distances. Still, because they transform trait observations into distances, they sacrifice the ability to discern differences among individual traits and obscure links between traits and mechanisms. Approaches based on trait spaces preserve information about the underlying biology by preserving the information captured in individual traits and trait–environment relations, but may be computationally difficult at higher dimensions. Within the trait space view, we often focus on quantifying a single trait dimension at a time because we have theoretical reasons for focusing on particular traits (which may reflect trait syndromes; e.g. the leaf economic spectrum). However, we note that ecological and evolutionary forces will operate on whole phenotypes, and focusing solely on independent trait axes may miss important dynamics caused by trait covariances (Laughlin & Messier, 2015; Pistón et al., 2019). While we touch on all of these approaches in this work, our focus is primarily on those that characterize the shapes of a single dimension of the multidimensional trait space at a time.

The overall shapes of distributions can be characterized by their means and higher statistical moments—the variance, skewness, and kurtosis (Box 1). These moments of trait distributions are a key component of Trait Driver Theory (Box 1; Supporting

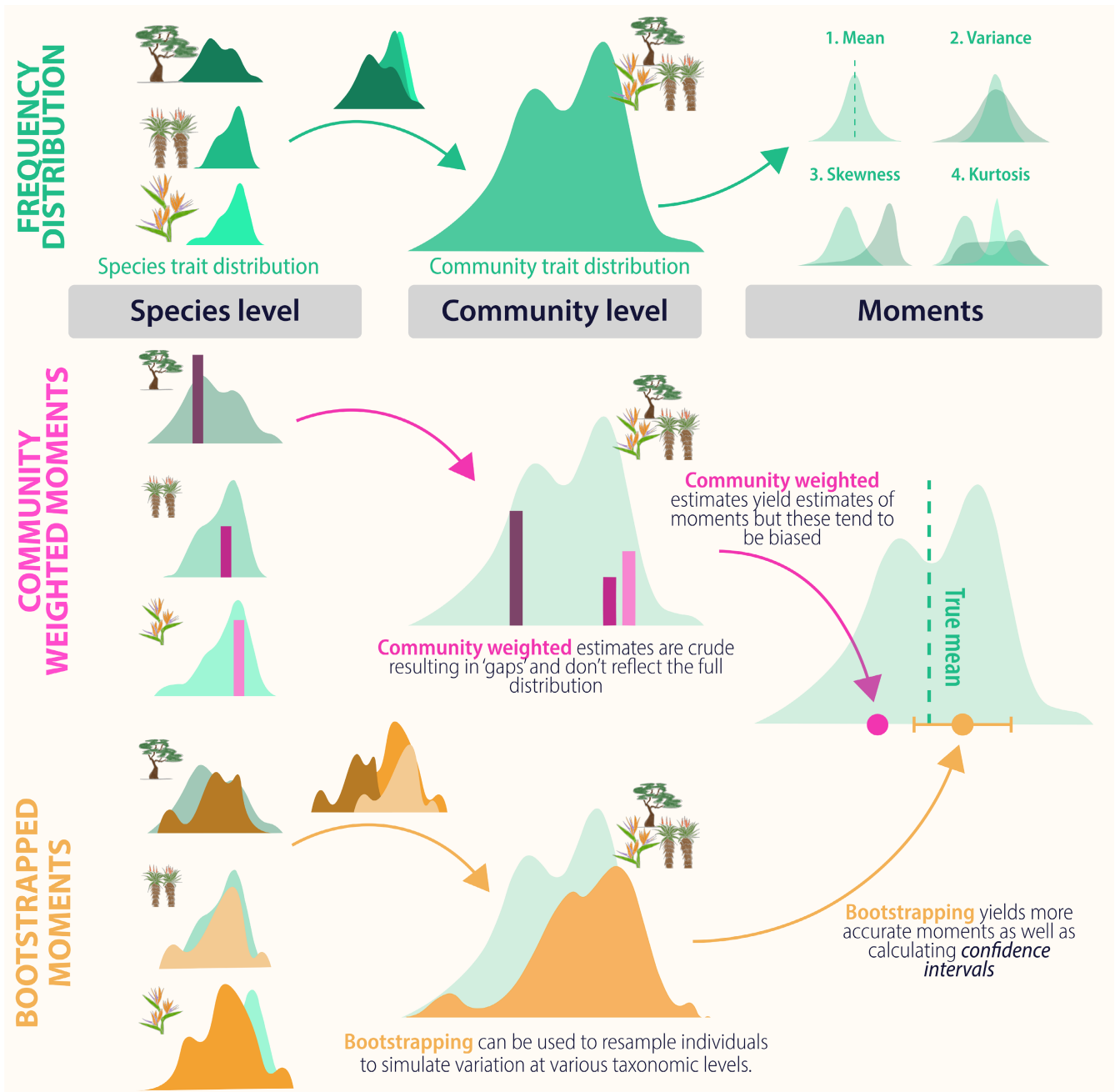


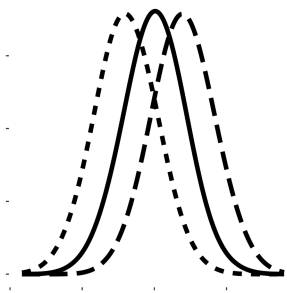
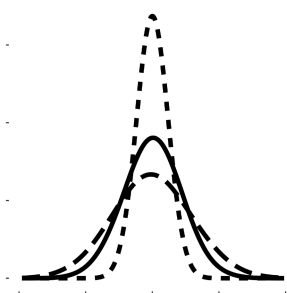
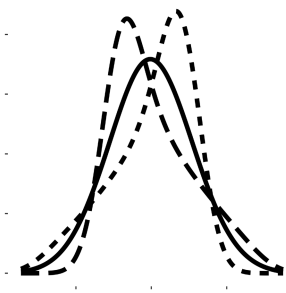
FIGURE 1 Inferring community trait distributions. *Top row:* An example community of three species in the Johannesburg Zoo (JBZ). These species can all be characterized by a common trait (as they are all plants, say, leaf size), and there is variation among and within species and individuals in the value of that trait. Thus, the JBZ population of each species can be characterized by a distribution of trait values, and the combination of these trait values across all individuals of all species yields the JBZ community trait distribution. The first raw moment is the mean. It describes the central tendency of the distribution. The second central moment is the variance, which describes the spread of the data. The third standardized moment is the skew, which measures the lopsidedness of the distribution; symmetric distributions will have a skewness, if defined, of zero. The fourth standardized moment is kurtosis, which measures the heaviness of the distribution's tails. High kurtosis corresponds to a distribution with long tails compared to the normal distribution, while low kurtosis corresponds to relatively short tails. As the kurtosis of normal distribution has a value of three, it is common practice to instead calculate the 'excess kurtosis' by subtracting three from the kurtosis. This excess kurtosis is sometimes simply called 'kurtosis', a convention we follow here. The mean and the central and standardized moments describe the shape of the trait distribution and can be used to infer ecological processes and functions. *Middle Row:* Since it is usually not feasible to measure all individuals (or all organs, e.g. leaves, teeth, etc.), a common approach is to measure the traits of a subset of individuals and represent each species by its mean trait value. The traits of these species are then combined with some measure of species abundance (e.g. coverage, body size, biomass) to estimate the mean and additional moments of the community trait distribution. *Bottom row:* Rather than collapsing the variation across all the measured individuals into a single mean value per species, we can employ a bootstrapping approach that uses all the measurement variations. With this approach, we randomly resample (with replacement) the observed trait distributions (or statistical distributions fitted to the data) in proportion to some measure of abundance. This method utilizes the same data but, as we demonstrate below, provides results closer to the true values and provides confidence intervals around the estimated values.

Information [Appendix 1](#); Enquist et al., 2015; Maitner et al., 2017; Weiher & Keddy, 1995; Wiczyński et al., 2019). The shape of the trait distribution can reveal information about the processes shaping and affecting population and community composition and species abundance, which will, in turn, impact ecological processes and ecosystem responses ([Box 1](#); Supporting Information [Appendix 1](#); Enquist et al., 2015; Weiher & Keddy, 1995). This information can rule out potential mechanisms and support

others in observational or correlative studies. We can pair it with experimental approaches to disentangle alternative mechanistic hypotheses explicitly and thus link trait distributions with underlying processes (Weber & Agrawal, 2012). While there are challenges to inferring community dynamics from trait diversity patterns, several potential solutions have been suggested, including incorporating mechanistic models and temporal data (Münkemüller et al., 2020).

BOX 1 Why measure the moments of trait distributions?

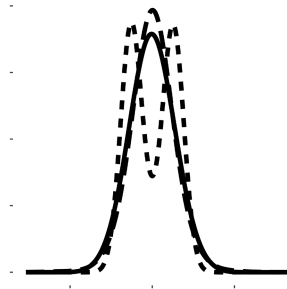
Each trait moment is associated with specific ecological processes. Different measures of the trait distribution can provide insights regarding the main drivers of species composition as well as the past and future dynamics of the species assemblage and can have consequences for ecosystem functioning.

Moments of trait distribution	Ecological processes	Ecosystem response
<p>Mean</p> 	<ul style="list-style-type: none"> • May represent local phenotypic optima (Muscarella & Uriarte, 2016; Todd et al., 2011) • Dominant phenotype due to competitive hierarchy (Freckleton & Watkinson, 2001) 	<ul style="list-style-type: none"> • Will shift productivity according to form of growth equation (Enquist et al., 2015) • Shifts in mean body size can cause trophic cascades (Shackell et al., 2010)
<p>Variance</p> 	<ul style="list-style-type: none"> • Low where abiotic filtering is strong (Keddy, 1992; Weiher & Keddy, 1999) or competitive exclusion is high • High where there is high immigration, competitive niche displacement and/or temporal variation • High where there are repeated disturbances or environmental variability (Grime, 2006) • High where species pools overlap (Denelle et al., 2019) 	<ul style="list-style-type: none"> • Increased variance accelerates community response to environmental changes (Enquist et al., 2015; Norberg et al., 2001) • Increased variance can reduce ecosystem productivity and fluxes (Enquist et al., 2015; Norberg et al., 2001) • Increased variance in one trophic level can promote increased variance in interacting trophic levels • Higher variance may increase intraguild predation (Wissinger, 1992) • Variance can stabilize or destabilize species coexistence (Bolnick et al., 2011; Hart et al., 2016; Uriarte & Menge, 2018)
<p>Skewness</p> 	<ul style="list-style-type: none"> • Strong skewness can indicate a response to rapid environmental changes (Enquist et al., 2015). • Weak skewness may result from species introductions (Blanchet et al., 2010) • Low skewness may reflect the loss of traits in distribution tails due to filtering, exclusion or ontogeny (Enquist et al., 2015) • High skewness may be caused by biased immigration (Enquist et al., 2015) 	<ul style="list-style-type: none"> • Changes in skewness will impact dependent trophic levels (Inouye et al., 2019) • Asymmetric loss of traits may cause a disproportionate decline in ecosystem functions (Duffy, 2003) • Skewness may impact ecosystem multifunctionality (Le Bagousse-Pinguet et al., 2021)

BOX 1 (Continued)

Moments of trait distribution

Kurtosis



Skewness + Kurtosis



Ecological processes

Here, TraitStrap calculates the **excess kurtosis** (see text). Excess kurtosis measures the 'tailedness' of distributions relative to the normal distribution. Distributions with negative excess kurtosis (platykurtic) are characterized by fewer and/or less extreme outliers than the normal distribution (more peaked distribution than the normal distribution). Distributions with a positive excess kurtosis (leptokurtic) have longer tails (more outliers) than the normal distribution

Positive kurtosis—A higher kurtosis corresponds to more outliers or a greater extremity of deviations

- Higher excess kurtosis—a more platykurtic (peaked) distribution—reflects competitive exclusion or other types of biotic exclusion or competitive hierarchies (Abrams & Chen, 2002; Navas & Violle, 2009), and can reflect a low trait diversity, and may typically occur under strong environmental filtering that would select for a limited range of trait values (Keddy, 1992).
- Lower excess kurtosis may result from species introductions (Blanchet et al., 2010) and can reflect the coexistence of functionally contrasting species (Enquist et al., 2017)
- Asymmetrical distributions ($\text{skewness}^2 > 0$) and lower kurtosis values than expected by chance imply heavier tailed distributions, with a higher relative abundance of rare species with extreme trait values. This pattern is consistent with trait differences between rare and common species promoting coexistence.
- For symmetrical distributions ($\text{skewness}^2 = 0$), observed kurtoses were also lower than expected by chance, consistent with high levels of trait diversity within dryland communities and co-dominance of functionally contrasting species (see Gross et al., 2017)

Ecosystem response

- Increased kurtosis in dispersal kernels causes increased rates of spread (Kot et al., 1996)
- Decreased kurtosis values reflect increased trait diversity (Gross et al., 2017), which may lead to increased productivity (e.g. Spehn et al., 2005)
- Functionally rare species (those in the tails of the distribution) may play important roles in maintaining ecosystem services (Delalandre et al., 2022)
- Indicator of multifunctionality (Gross et al., 2017)

1.2 | Community weightings are imprecise estimates of trait distributions

The current standard for summarizing phenotypic distribution relies on species average trait values and measures of species abundance (note: throughout, we reserve the word 'mean' for referencing the first moment of trait distributions and use 'average' in the broader sense; and we use 'abundance' to broadly represent different

quantifications of relative abundance in the assemblage, including biomass, coverage, number of individuals, etc.). The approach calculates the weighted moments of trait distributions by taking the average trait values for each species in the community and multiplying them by a measure of abundance (e.g. cover, biomass; Grime, 1998; see Supporting Information Appendix 2 for details). New measures and indices of trait diversity and multifunctionality rely on accurate measures of the four moments of trait distribution (e.g. the variance,

skewness and the kurtosis of the trait–abundance distributions; Gross et al., 2017).

The problem with community-weighted moments is that they represent the traits of a species with a single value, usually a species-level average. Attributing a property of a group (i.e. species) to all individuals within that group leads to an aggregation bias termed the ecological fallacy (Clark & Avery, 2010; Denny, 2017; Ruel & Ayres, 1999). For example, imagine we sample two sites along an environmental gradient, and the individuals at these sites are roughly the same body size. However, many of the individuals at one of these sites belong to species that typically grow much larger. Based on individual measurements, we would correctly conclude that there is no relationship between the gradient and individual size. In contrast, if we use species-average traits instead of locally measured traits, we may erroneously infer a significant relationship between size and the environmental gradient. The ecological fallacy is caused by variation within the groups we aggregate by, which in practice are often species. Traits vary within species due to ontogeny, abiotic forces, biotic interactions, experimental treatments and genetic differences, and thus, average traits may be a poor approximation for the traits of an individual (Agrawal, 2001; Bolnick et al., 2011; Finney et al., 2002; Hendry, 2016; Miner et al., 2005; Parmesan & Yohe, 2003; Reznick & Ghalambor, 2001; Violle et al., 2012). Population biologists and ecologists have, therefore, increasingly come to realize that locally measured traits, along with the representation of intraspecific trait variation, are essential and can better reveal the local ecological and evolutionary processes that shape local communities (Bolnick et al., 2003, 2011; Fontana et al., 2018; Hart et al., 2016; Jónsdóttir et al., 2022; Lake & Ostling, 2009; Siefert et al., 2015; Uriarte & Menge, 2018; Violle et al., 2012). Furthermore, using average species traits leads to statistical non-independence between sites that share species, inflating Type 1 errors when using standard regression techniques (Miller et al., 2018). An alternative approach is to use traits that are measured and averaged within a smaller scale (e.g. site, plot). Using site-specific average traits incorporates intraspecific variation across sites. It alleviates the issues of inflated Type 1 error. Still, it fails to capture variation within sites and does not account for uncertainty (i.e. it does not provide an estimate of error). The problem of aggregation bias remains regardless of the aggregation scale, although the magnitude of the problem may decrease towards finer sampling resolutions.

Accurately assessing the shapes of the realized trait distributions of populations and communities depends on sampling and estimation methodologies. If we consider a set of traits (e.g. all plant heights in a 5-m plot, all fish biomasses in a 100-L mesocosm, leaf area of all sun-exposed leaves in a 1-ha plot) as a statistical population with some parameter we would like to quantify (e.g. mean, variance), we usually try to represent that population using a statistical sample. To accurately depict the population, that sample must be representative. The estimator used should be consistent (approach the true value as sample size increases), relatively efficient (have a small variance relative to other estimators) and unbiased (expected value of the estimator is the true value; Everitt & Skrondal, 2010). Previous studies

have examined how different sampling methods and estimators perform when estimating trait mean and variance (Baraloto et al., 2010; Carmona et al., 2015; Lavorel et al., 2007; Paine et al., 2015), and suggest that community-weighted approaches can provide reasonable estimates of mean and variance. However, this previous work was limited to community-weighted approaches, did not include skewness or kurtosis and focused on a single taxon in each study.

1.3 | Bootstrapping incorporates variation to improve estimates

Community-weighted moments ignore intraspecific variability, leading to biased estimates of trait distributions and implicitly misrepresenting uncertainty (Figure 1, middle row). To account for intraspecific variation, we need to move from quantifying species as a single, average trait value to a distribution of trait values (Carmona et al., 2016). Previous efforts to integrate intraspecific variation into measures of trait diversity have done so by representing species as multiple individuals (Cianciaruso et al., 2009) or full distributions (Carmona et al., 2016; Mason et al., 2008; Mouillot et al., 2005; Wong & Carmona, 2021). While these efforts help solve the problem of aggregation bias, they do not fully address the issue of uncertainty. However, we can account for intraspecific variation and uncertainty through the statistical sampling of the raw data or fitted distributions to better estimate interspecific and intraspecific variation in traits (Enquist et al., 2015). Bootstrapping is a method that uses random sampling with replacement to generate a set of distributions (Figure 1, bottom row). Moments of the distribution are then estimated by averaging the moments across the set of distributions, and their uncertainty can be reported by summarizing their variation within the set (Davison & Hinkley, 1997; Efron, 1979). *Nonparametric bootstrapping* involves random sampling with replacement from the full set of observed data to generate a set of new distributions. In the case of trait distributions, the observed traits are sampled in proportion to their abundance (e.g. the relative abundance of a species within a community and trait values within each species).

Nonparametric statistical procedures rely on no or few assumptions about the shape or parameters of the population distribution from which the sample was drawn. Thus, they may be beneficial when the underlying distribution is unknown or is not easily fit by a parametric distribution. *Parametric bootstrapping* assumes that the data come from a known distribution but with unknown parameters (e.g. a normal distribution). The method then estimates the parameters from the sampled trait data and uses the estimated distributions to simulate samples. The assumption is that the sample set of trait data comes from a population characterized by a probability distribution that a fixed set of parameters can adequately model. Parametric approaches are useful if we have a priori reasons to suspect a certain probability distribution. Although our focus is predominantly on statistical moments of univariate distributions, bootstrapping approaches can be used

to account for uncertainty in any quantification of trait structure, including multivariate metrics or entire multivariate distributions. Due to their generality, bootstrapping methods can easily be integrated into existing trait diversity frameworks (e.g. Carmona et al., 2016; Cianciaruso et al., 2009). However, whether using a parametric or nonparametric bootstrapping approach or quantifying a univariate or multivariate metric, bootstrapping leverages the variation within the data to estimate trait distributions better and provide uncertainty estimates (see Supporting Information Appendix 3). Bootstrapping also encourages us to conceptualize the traits of a population or sample as distributions rather than point estimates (Figure 1) and, in doing so, reflects issues inherent to many standard sampling protocols.

1.4 | Common sampling protocols may lead to biases in trait data

Accurate characterization of trait distributions requires representative sampling of the traits of a particular statistical population. For example, suppose we are interested in a particular plot's community dynamics or ecosystem function. In that case, the true statistical population of interest will be the complete set of values of our focal traits within that plot (e.g. the height of every plant or the area of every leaf). Conversely, for other questions (e.g. trait evolution), we may want to standardize measurements by limiting our statistical population to a particular subset (e.g. adult body mass, sun-exposed leaves). Common sampling protocols in trait-based ecology (e.g. Pérez-Harguindeguy et al., 2013) often make choices that limit accurate measurement of the full statistical population, for example, by focusing the collection on 'fully grown' (i.e. large) and healthy individuals or organs, with little attention to representing the complete variation within the local population. Even where size bias is unintentional, larger individuals are more likely to be measured because they are easier to identify, see, catch or measure in macroorganisms, or conversely, some methods are more likely to sample smaller individuals (e.g. in microbial ecology) or those in poor condition (Bisi et al., 2011; Boonstra & Rodd, 1982).

Another common practice when measuring traits in ecology is to limit field data collection to only the most common species. A common practice is collecting field data from the most abundant species until 80% of the community's individuals are represented (Pérez-Harguindeguy et al., 2013), automatically excluding rare species. This practice may also bias estimates of trait distributions, particularly when the trait values of rare species are outliers (Rosenzweig & Lomolino, 1997; Violle et al., 2017). It is also common practice to forego local trait measurements altogether and instead use publicly available trait values for the species of interest. Trait sampling is time-consuming, and local trait datasets from the study system are often incomplete or unavailable. Therefore, trait values are often taken from sources that aggregate or summarize global trait data (e.g. Bjorkman et al., 2018; Jones et al., 2009; Maitner et al., 2017; Oliveira et al., 2017), sometimes using inferred traits (e.g. genus-level average

traits or traits inferred via statistical modelling) where species-level measurements are lacking (Penone et al., 2014; Swenson, 2014). A final common approach is to randomly sample a minimum number of individuals from within the community, irrespective of species identity (Baraloto et al., 2010). This ensures that intraspecific variability is representatively captured within all species. However, given the long tails of species abundance distributions (Enquist et al., 2019), purely random sampling may overlook rare species unless sample sizes are quite high. Each of these common sampling protocols may be well suited for particular research questions or settings; however, if they are used to estimate the trait distribution of a full statistical population rather than a particular subset (e.g. the realized trait distribution of a community, aiming to understand the functioning of that community), they may bias estimates.

Although bootstrapping methods are becoming increasingly common in the trait-based literature (Enquist et al., 2017; Gaüzère et al., 2019; Henn et al., 2018; Jónsdóttir et al., 2022; Ross et al., 2017), we currently lack a clear picture of the relative performance of these methods and how much common sampling strategies impact those performances. Here, we test the accuracy of different methods to estimate the four moments of trait distributions. To do so, we developed an R package (R Core Team, 2020), *TRAITSTRAP*, that estimates the moments of community trait distributions or alternative relevant metrics used in trait-based ecology along with associated uncertainty using parametric and nonparametric bootstrapping methods. The package offers flexibility by enabling multiple, user-specified sources of traits and trait variation and hierarchical sampling designs, and may hence be adapted to cope with variation in data availability as well as the needs of different specific research questions. The R package estimates trait distributions and can also be used to calculate the community-weighted moments of trait distributions, along with confidence intervals and uncertainty estimates. Using *traitstrap*, we analysed various datasets across different organisms and ecosystems to test how data collection strategies and methods of estimating the moments of community trait distribution impact their reliability.

2 | MATERIALS AND METHODS

2.1 | The *TRAITSTRAP* Package

Our R (R Core Team, 2020) package, *TRAITSTRAP* (<https://github.com/Plant-Functional-Trait-Course/traitstrap>), estimates the moments of community trait distributions using bootstrapping approaches. The package is available via CRAN (<https://cran.r-project.org/package=traitstrap>), and can be installed using the R function `install.packages("traitstrap")`. Two inputs are needed: a dataset with information on the abundances of species within a community (used to weight species traits); and a dataset with the traits and trait values for the species in the community. *Traitstrap* is based on tidyverse (Wickham et al., 2019) code, and the data should be organized in a tidy long format. The two datasets need some specific columns (i.e.

taxon, abundance, trait and trait value), but the naming convention is flexible and specified in the functions. More details on the package are included in Supporting Information Appendix 4. Example workflows can be found in Supporting Information Appendix 5. Example data, code and workflows are also included in the traitstrap package and can be viewed using the R function vignette(package = "traitstrap").

2.2 | Comparing method performance on five empirical datasets

We used five datasets (Table S1, Appendix 7) to evaluate the performances of two different bootstrapping (BS) approaches (parametric and nonparametric) against two approaches for calculating the classical community-weighted (CW) moments (using cross-site vs. site-specific species average traits) across a variety of sampling designs. We sought diverse datasets so that differences observed between methods were as generalizable as possible. These datasets (Table S1) included both plants and animals from aquatic and terrestrial systems that vary in richness and abundance and included both observational and experimental work. We treat each dataset's full set of trait observations as the 'true' trait distribution for the community and compare the inferred community-weighted and bootstrapped moments with these true values (Baraloto et al., 2010; Carmona et al., 2015; Lavorel et al., 2007; Paine et al., 2015).

We acknowledge that the ability to which our 'true' distributions reflect reality will differ across datasets. For the herbaceous plants' data, the individual-level traits were measured for every individual within each plot and thus *do* reflect the full trait distributions for that study area. Conversely, other datasets may not sample every individual but include wider spatial or temporal coverage. While we conduct analyses for multiple datasets where possible, we place special emphasis on the herb dataset, because (1) it contains complete trait distributions for individual-level traits within spatially delimited communities (and one measurement per individual for organ-level traits); and (2) it contains very high levels of intraspecific variation, making it a challenge for our methods. Using these five datasets and four approaches to estimating the moments of community trait distributions, we conducted simulations comparing method performance across a gradient of sampling intensities. For subsets of the datasets, we also compared the performance of the methods in the presence of size-biased sampling, abundance-biased sampling and when using local traits versus those available in a global trait

database. Full methodological details are available in Supporting Information Appendix 8.

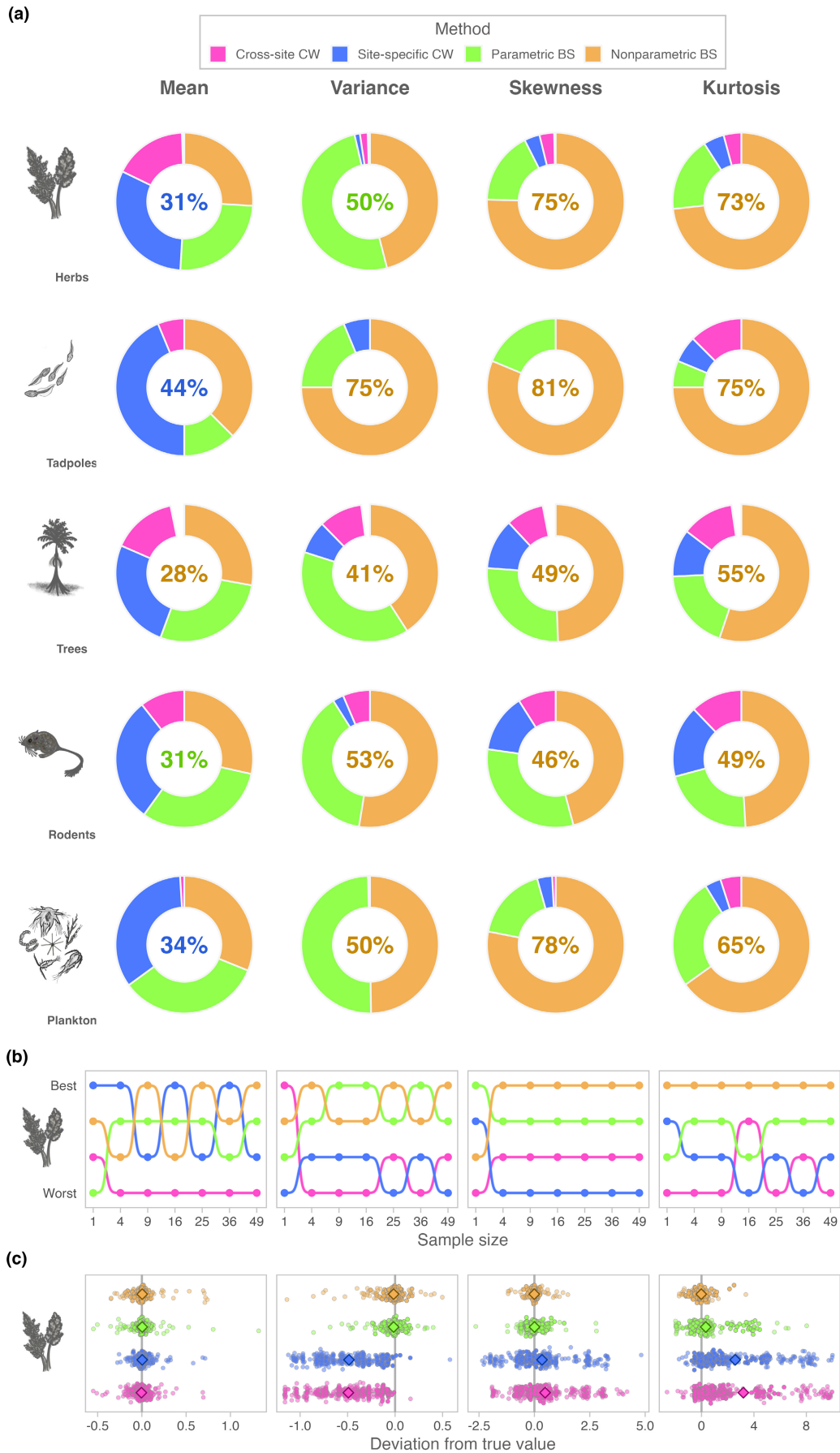
To evaluate the performance of bootstrapping approaches in calculating distance-based metrics, we compared the performance of parametric bootstrapping against community-weighted (using cross-site vs. site-specific species average traits) methods for calculating five common multivariate metrics: functional evenness (FEve; Mason et al., 2005), functional dispersion (FDis; Laliberté & Legendre, 2010), Rao's Quadratic entropy (RaoQ; Botta-Dukát, 2005), functional richness (FRic; Villéger et al., 2008) and functional divergence (FDiv; Villéger et al., 2008). We additionally calculate two metrics of functional rarity: average functional uniqueness and average functional distinctiveness (Grenié et al., 2017; Violle et al., 2017). Using the herb dataset, we simulated a gradient of sampling intensities. We evaluated the performance of the three methods relative to the 'true' FEve, FDis, RaoQ, FRic, FDiv, average uniqueness and average distinctiveness values. Full methodological details are available in Supporting Information Appendix 8. We expect bootstrapping to perform well for the metrics FDis, RaoQ and average distinctiveness, as these metrics are functions of the full trait-distance matrix. However, FEve is a measure of point regularity along a minimum spanning tree in trait space, FRic and FDiv are both sensitive to the borders of trait space, and functional uniqueness is a measure of the minimum distance between taxa. Thus, we expect these metrics to be biased by the resampling procedure that underlies bootstrapping. Still, we include these metrics to highlight the limitations of bootstrapping.

3 | RESULTS AND DISCUSSION

3.1 | Nonparametric bootstrapping: Simply the best

Across datasets, bootstrapping outperforms community-weighted calculations of trait moments in terms of both bias and precision (Figure 2c). These methods consistently had lower errors and confidence intervals that generally included the true moment values (Figure 2). While all methods perform reasonably well for estimating means (Figures 2c–4), differences in performance between the bootstrap and community-weighted approaches are apparent for estimating the higher moments (Figure 2c). For example, across all datasets, at sample sizes of nine individuals per species per plot, nonparametric bootstrapping reduced the percent error by an average of 26% for the mean, 74% for the variance, 73% for the skewness and 67% for the kurtosis, relative to cross-site community-weighted

FIGURE 2 Method performance across moments and dataset. (a) Best-performing methods (closest to actual values and containing the true values within their CI) across simulations, arranged by dataset (rows) and central statistical moments (columns). Situations where no method included the true value in its CI were not counted; hence, some donuts are incomplete. The 'winning' method (and the percentage of times it was the winner) for each trait x moment combo is also shown. (b) Different methods ranked by lowest mean deviation ('winner' is the top row) for sample sizes from 1 to 49 trait measurements per species. Mean deviation was calculated only for values with the true value in their CI. (c) Points denote estimated values of central statistical moments with vertical lines denoting the true value of each moment and diamonds denote the mean estimated deviation for each dataset-moment method.



estimates (Table S3). These general patterns hold across various sampling intensities and the full datasets (Figure 4b, Figure S4). Trait moments estimated via community-weighted approaches showed relatively weak relationships with the true moments (Figure 3), suggesting that their use may impede our ability to correctly infer the drivers of trait distributions (e.g. trait by environment relationships) or their impacts (e.g. ecosystem function). We also observe some consistent biases in the directionality of estimates of the different moments, where both community-weighted methods consistently underestimate the variance and overestimate skewness and kurtosis relative to the true moments (Figure 2c, see Figure S5 for all datasets and Figure S6 for comparisons across traits). Biases were observed across organisms and traits. These biases occur because the moments beyond the mean are sensitive to the shape of the distribution, and by ignoring trait variation within a species, we are left with a poor estimate of that shape.

While the differences between the two bootstrapping approaches are more subtle (Figure 2c), nonparametric bootstrapping fairly consistently ranks above the parametric approach (Figure 2a,b), and it is also less likely to introduce systematic biases in the different moments (Figure 4). In addition, nonparametric bootstrapping is not only consistently the better method across the different moments and datasets (i.e. organisms and dataset types) but also at the trait

level, including both chemical and morphological traits (e.g. leaf carbon content vs. leaf area, Figure S7) and individual- and organ-level traits (e.g. plant height vs. leaf size; Figure S8). Thus, nonparametric bootstrapping is the method most likely to yield the most accurate estimates and the least likely to introduce consistent biases.

Although our main focus here was on univariate trait distributions, our multivariate simulations likewise show that multivariate metrics were better estimated by bootstrapping than community-weighted approaches (Figure S16). Community-weighted approaches also consistently failed to correctly identify the relative order of sites (Figure S17). While FEve, FRic, FDIV and average uniqueness were better estimated by bootstrapping than either community-weighted approach, none of the approaches could correctly calculate either metric (Figure S16). Due to the difficulties of estimating FEve, FRic, FDIV or average uniqueness, we recommend caution when using these metrics.

3.2 | Even modest sampling yields relatively accurate estimates: That's all

Empirical trait datasets vary greatly in sampling intensity, but many studies have relatively moderate numbers of true replicate

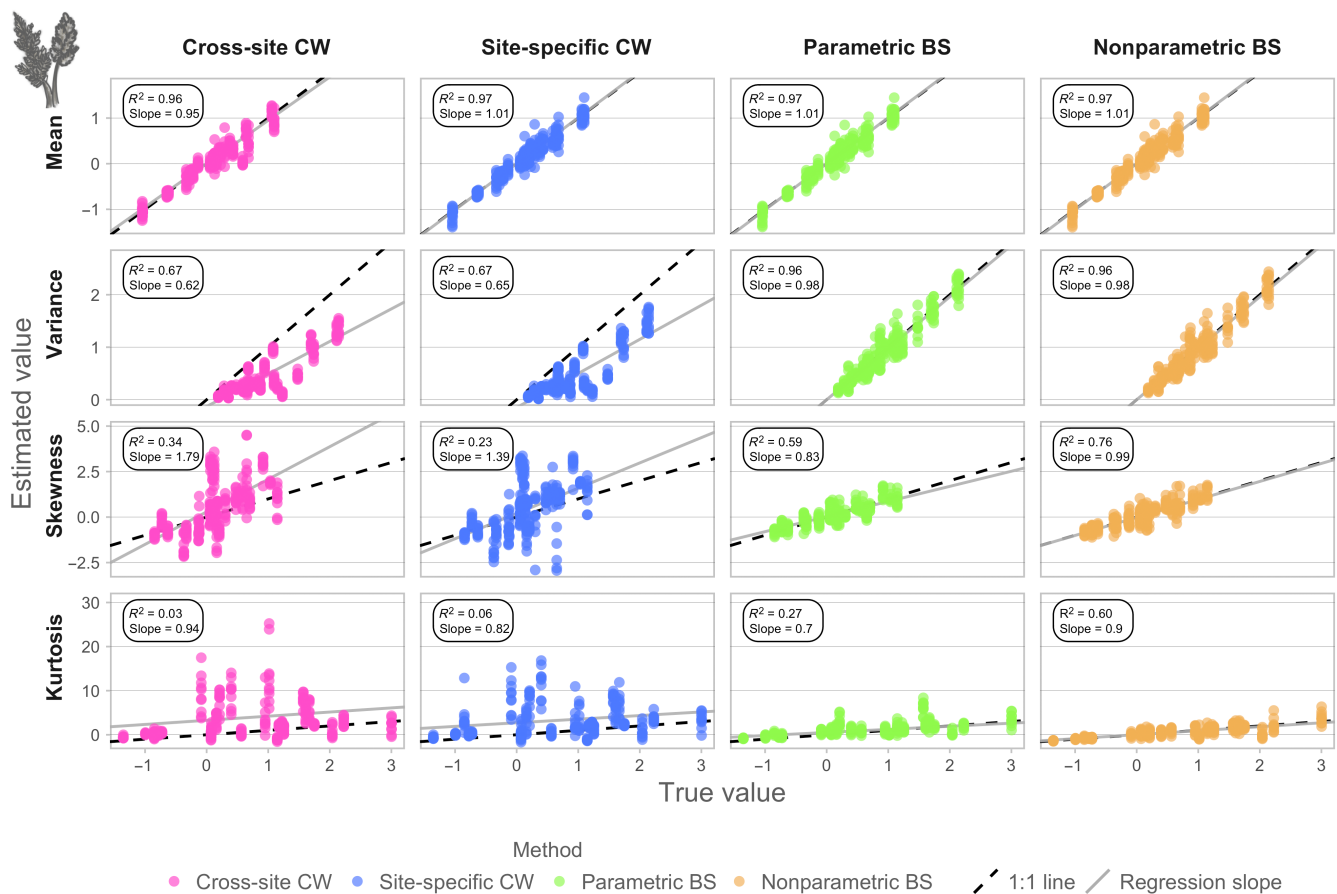


FIGURE 3 Correspondence between estimated and true values varies with method and moment. Points include 10 replicates across five traits and five sites. The sample size was restricted to a maximum of nine trait observations per species per site in these simulations, with fewer being sampled if there were fewer than nine individuals of a given species within a given site.

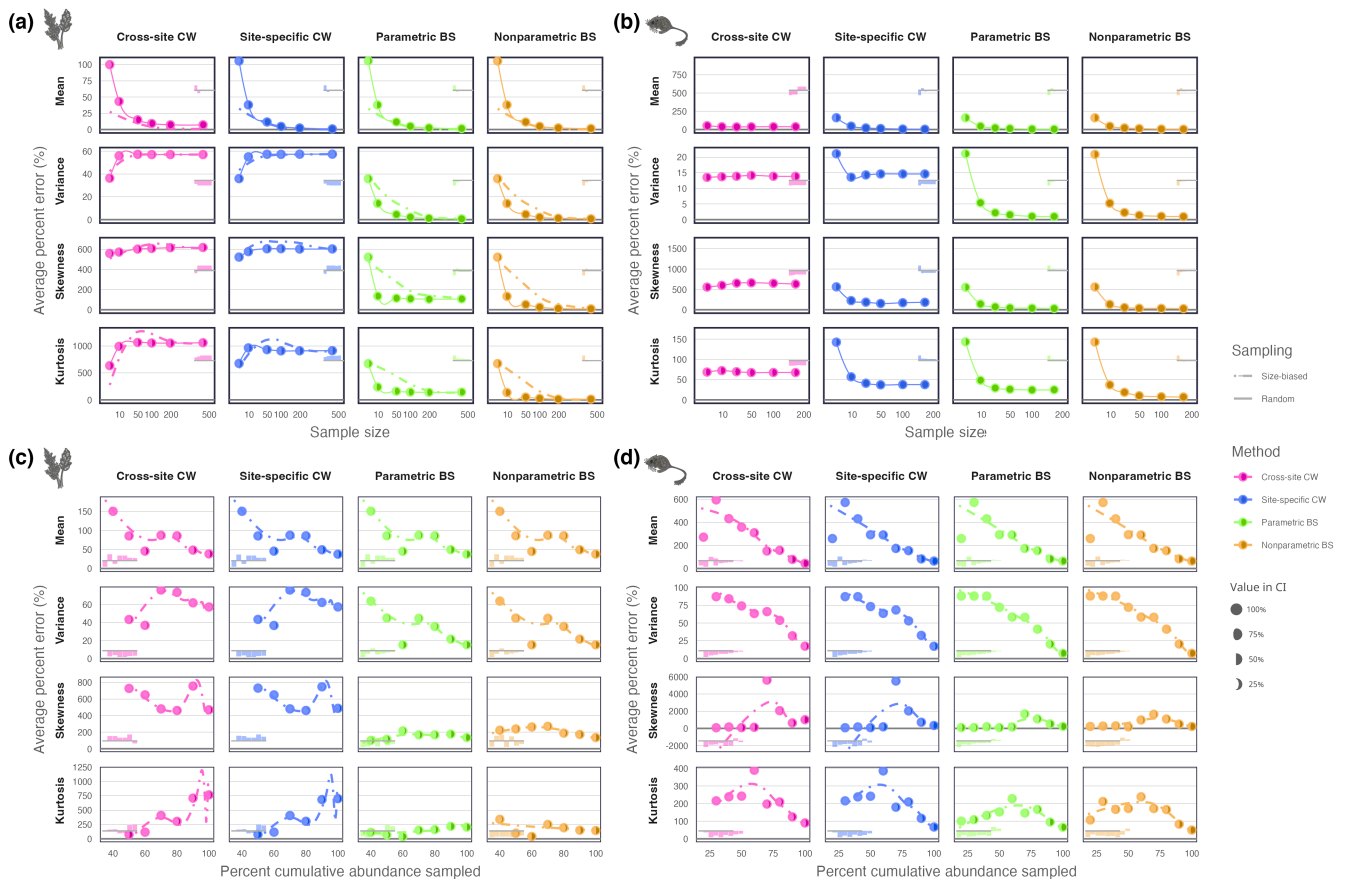


FIGURE 4 Effect of sample size and method on estimation accuracy. (a, d) Filled circles represent the average percent error between the estimated moment and true moment, averaged across sites and traits. The darker portion within each circle denotes the average percent of replicates for which the true mean fell within the 95% CI. Solid lines indicate random sampling, and dashed lines reflect sampling bias towards larger leaves (herbs) or individuals (mammals). The sample size represents the maximum number of individuals per species, with fewer being sampled if they were not in the community. Inset figures show tendencies in bias (over- vs. under-estimation). (a, b) Dashed line represents biased sampling (c, d) sample sizes were fixed at nine randomly sampled individuals.

trait measurements (i.e. number of measurements per sampling unit, which could be population, treatment, site etc. within species). To assess the consequences of sampling effort for trait moment estimation and to explore any potential differences between methods and moments in coping with low sampling effort, we used simulations where we varied the number of individuals for each replicate sampling unit from one individual to all individuals. Overall, moment estimates tend to asymptote around a sample size of ~ 9 trait values, suggesting that relatively modest sampling intensity gives stable and reliable trait moment estimates (Figure 4). At this sampling intensity, bootstrapped estimates almost always contain the true value in their CI. The degree to which the estimated moments of community trait distributions improve with increasing sampling intensity varies across traits, datasets, moments and methods (Figure 6, Figures S7 and S8). In contrast, community-weighted approaches often fail to yield accurate estimates, especially for the higher moments of trait distributions, even after sampling all trait values in the community (e.g. cross-site community-weighted kurtosis for the herbs dataset; Figure 4). Importantly, the relationship between the true

and estimated moments weakened as we went from the first moment (mean) to the fourth (kurtosis). Still, weakening was much more pronounced in community-weighted approaches than in bootstrapping approaches (Figure 3). Bootstrapped estimates of multidimensional metrics similarly outperformed community-weighted approaches, nearing the asymptote at sample sizes of ~ 9 for metrics that could be accurately estimated (i.e. FDis, RaoQ, average distinctiveness; Figure S16). The fact that bootstrapping reaches its 'optimal' asymptote at a sample size of approximately 9 (as well as having the estimate that is closest to the true value, Figure 4) is promising for data collection as this implies that even at a relatively modest sampling effort, we are still able to make reasonably accurate estimates for the moments. We note that bootstrapping can be applied even with small sample sizes (e.g. 30 total samples; Chernick, 2011). At sample sizes of one trait value per species, bootstrapped estimates will equal community-weighted estimates, making bootstrapping a viable option even when data are limited.

Unfortunately, the aggregation biases that cause community-weighted methods to perform relatively poorly when using locally

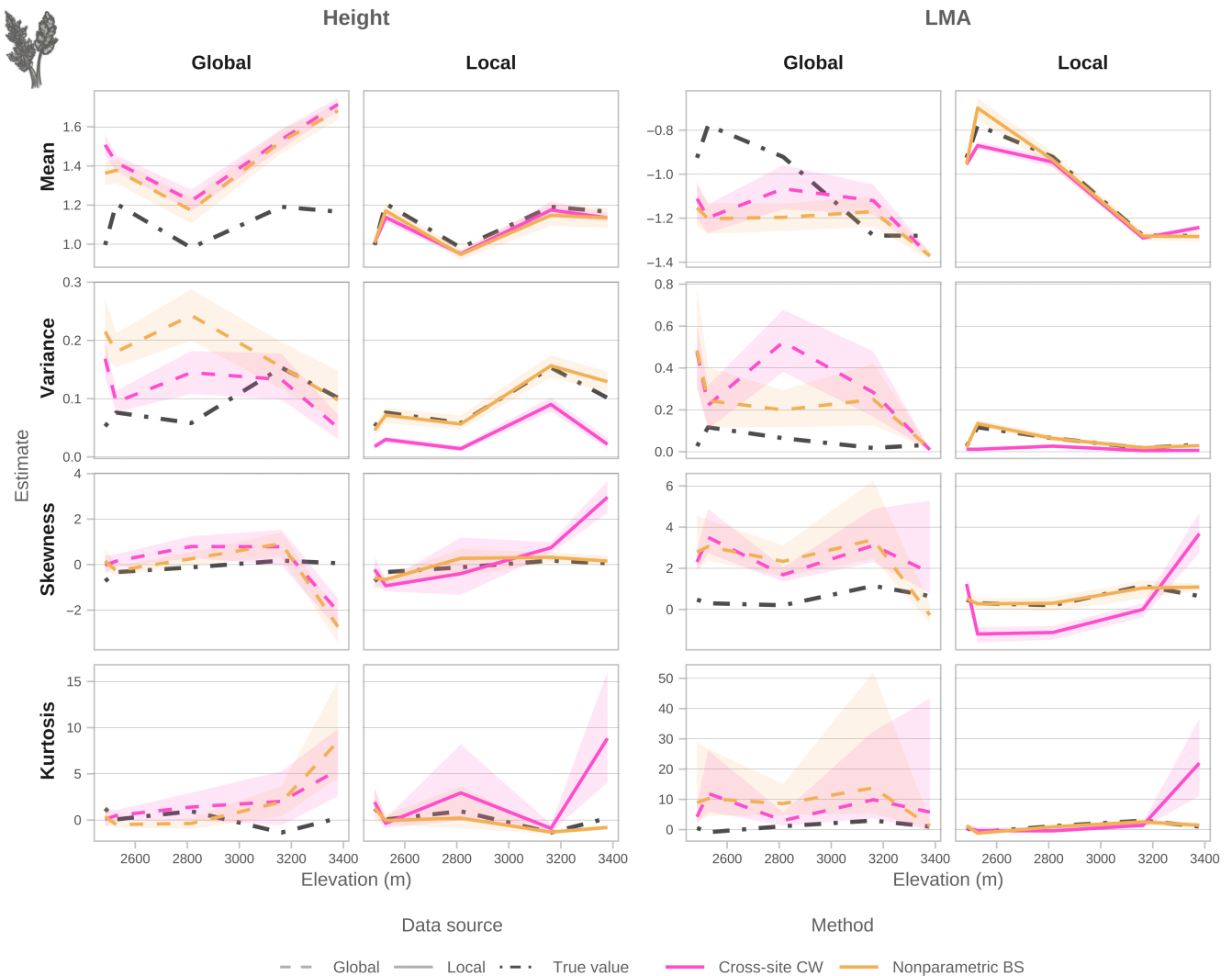


FIGURE 5 Effect of locally sampled traits versus traits from a global database on estimating the shape of trait distributions. Non-parametric bootstrapping best captures shifts in trait distributions across gradients. For example, two traits show that when traits are sampled from global databases (instead of measured locally incorporating intraspecific trait variation), the moments of trait distributions tend to be overestimated and tend to provide an incorrect relationship with how traits shift across elevation. However, incorporating interspecific trait measures locally with intraspecific information combined with non-parametric bootstrapping provides a robust accurate measure of the shape of trait distributions across elevation. The dotted-and-dashed grey line denotes the true value of each moment. The solid-coloured lines represent an estimate of each moment calculated using a single random sample of (up to) nine individuals per taxon per site. The dotted, coloured lines represent an estimate of each moment calculated using data from a publicly available database. The shaded area denotes the average CI of the estimate. For clarity, we only present two estimation methods, but we note that the two bootstrapping methods performed similarly, as did the two community-weighting methods.

measured data will be exacerbated when using data collected across larger spatial or taxonomic scales, as we may rely upon when using global traits databases (Figure 5). Furthermore, applying bootstrapping approaches does not substantially improve the correspondence between estimates and the true values (Figure 5). However, we note that there may be ways to improve these estimates using global data, for example, by preferentially utilizing data from climatically similar conditions, prioritizing more closely related species or leveraging allometric relationships between measured traits. This highlights the importance of locally measured traits and a pressing need for work exploring how best to utilize global trait datasets where local sampling is impossible (e.g. at global scales or in the past).

3.3 | Biased sampling: Out of touch

3.3.1 | Size-biased sampling

Our simulations reveal that biased sampling of larger individuals within the community results in a corresponding bias in the estimates of moments (Figure 4). These size-biased estimates of moments can, to some extent, be compensated for with increased sampling intensity. However, this requires a considerable increase in sample size, from around nine in the unbiased case to more than 100 individuals per sampling unit when sampling is biased (Figure 4). Although we find that a larger sample size is required to reach the *same* 'optimal asymptote'

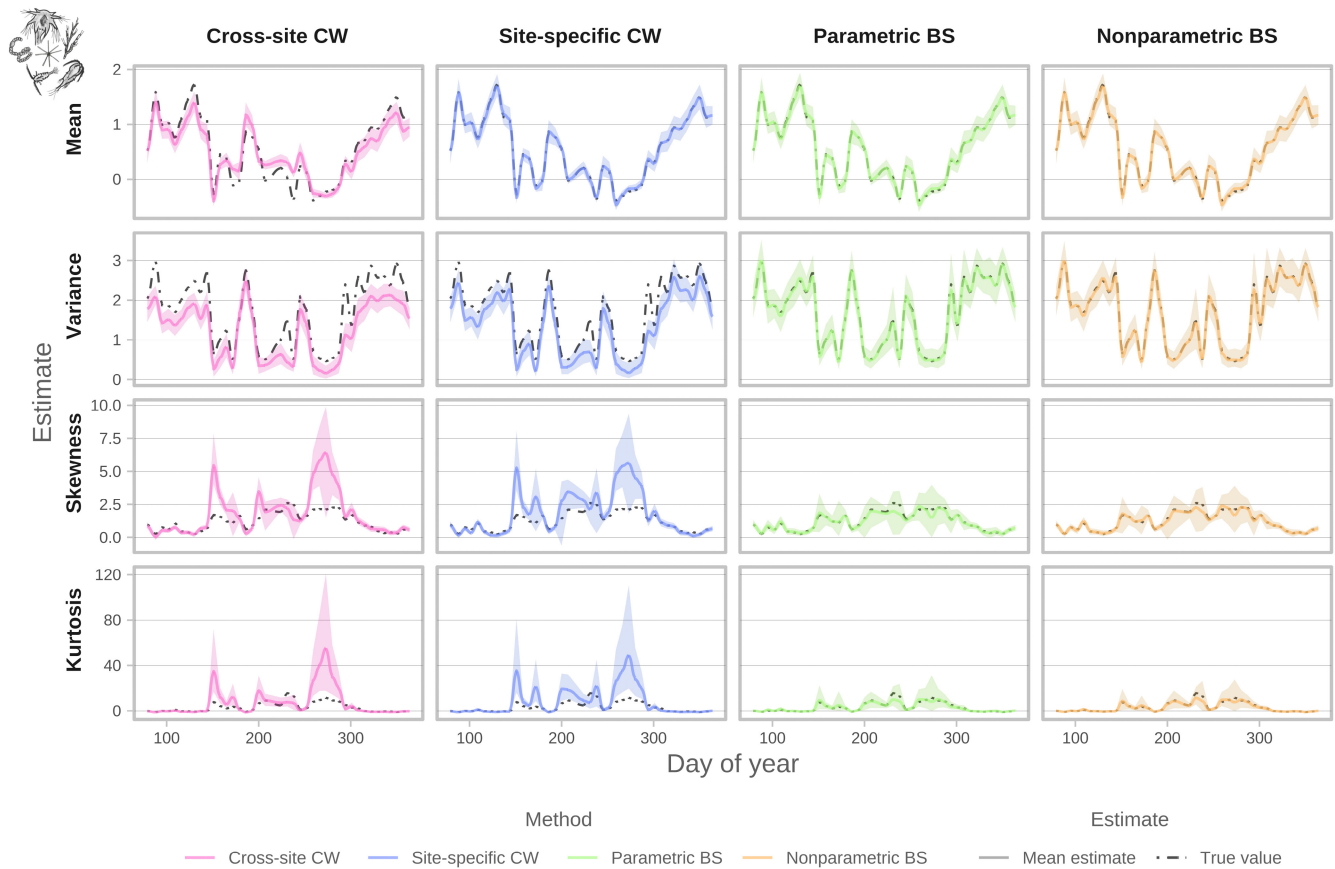


FIGURE 6 Temporal dynamics: changes in phenotypic moments over time. Nonparametric bootstrapping can accurately capture temporal shifts in trait distributions. Example showing plankton body size (log-transformed area) distribution changes from March 2019 to December 2019. The grey line denotes the true value of each moment. The solid-coloured lines represent the average estimate of each moment calculated using a sample size of nine individuals per taxon per day, calculated once every 7 days. The shaded area denotes the average CI of the estimate.

regarding deviation from the true value when compared to an evenly sampled community (dashed line vs. solid line in Figure 4), we observe bootstrapping still outperforms both community-weighted methods. This suggests that bootstrapping can control for size-biased sampling to some extent. However, given that sample sizes should be in the region of 100 individuals (an order of magnitude more than with unbiased sampling) to yield results comparable to random sampling, our findings suggest that the better, and more efficient, approach to obtain unbiased trait moment estimates is adjusting sampling procedures to ensure representative sampling of all species. However, we caution that bias is relative to the population of interest: What may be a biased sample for one study may be representative for another.

3.3.2 | Abundance-biased sampling

In simulations where sample size was fixed at nine individuals per species and we sampled species in decreasing order of abundance, we found that across all methods and moments, estimates only include the true mean in their CIs when we sampled close to 100% of species (Figure 4). When restricting the candidate species to only those that constitute 80% of the cumulative community abundance (as

is commonly recommended, e.g. Pérez-Harguindeguy et al., 2013), the estimates for *all* methods and moments have a greater deviation from the true value and rarely (if ever) contain it in their CI (Figure 4, Figure S11). In addition, it is often only after we have crossed the 80% threshold that estimates approach the true value—this is particularly clear in the rodents dataset for kurtosis estimates (Figure 4d). In simulations where sampling was species agnostic but was either (1) random concerning abundance; (2) biased towards common species; or (3) biased toward rare species, we found that random sampling performed the best overall (although we note that a bias towards rare species performed well for the rodent communities at a low sampling intensity; Figure S12). This strongly suggests that rare species are important in driving realized community trait distributions. Limiting sampling to only abundant species severely hampers our ability to confidently estimate the statistical moments of the realized trait distributions of entire communities.

3.4 | Uncertainty: Any way you want it

Traditional community-weighted approaches give point estimates, ignoring variation in the underlying data and uncertainty in the

estimates. In contrast, bootstrapping approaches embrace variation and uncertainty to improve estimation and provide confidence intervals for the resulting estimate. The confidence intervals generated by parametric and nonparametric bootstrapping perform well, even on our largest dataset, which included extremely variable temporal trends (Figure 6). Variations in uncertainty over space and time are also useful, as they may indicate differences in intraspecific variation or measurement accuracy. Our simulations of multivariate metrics show that bootstrapping can also improve the estimation of multidimensional metrics (Figures S14 and S15). Furthermore, bootstrapping approaches lend themselves to incorporating additional sources of uncertainty and propagating that uncertainty into downstream analyses (See Supporting Information Appendix 9). For example, the bootstrapping approaches discussed here can be readily integrated with evolutionary bootstrapping approaches (Felsenstein, 1988).

3.5 | Where to next: The promise

We compared the performance of different approaches for estimating the moments of trait distributions across a variety of sampling scenarios, taxa and datasets. We introduced the `TRAITSTRAP` R package to facilitate inferences of trait distributions via bootstrapping. We find that size-biased sampling protocols can yield comparable results to random sampling, but only at much larger (e.g. an order of magnitude) sampling intensity. Furthermore, sampling only the most abundant species, as advised in many trait protocols, prevents the accurate estimation of especially the higher moments of community trait distributions. Unfortunately, we also find that, even after accounting for uncertainty using bootstrapping, estimates of the moments of trait distributions using global trait data could not accurately capture local trends. This suggests that the general approach of randomly sampling a relatively modest number of individuals (9 individuals per sampling unit and species), focusing on covering all species in the community and analysing the data using nonparametric bootstrapping will enable reliable inference on trait distributions, including the central moments, of communities. We also compared the performances of different approaches for estimating multivariate metrics of trait distributions, finding that bootstrapping methods performed the best. However, we also found that some multivariate metrics could not be well estimated by any of the methods tested here. This suggests that these metrics may best be set aside until a reliable estimation method is found.

We make the following recommendations for future work:

1. Critically assess trait data needs to address your research question or hypothesis—what is your population of interest, and what scale, resolution and density of trait data will be required for a robust, unbiased assessment of the relevant parameters?
2. Explicitly include the major environmental drivers or other sources of key ecological or evolutionary variation relevant to the

research question or hypothesis in the trait sampling design (e.g. ecological gradients, experimental treatments).

3. Strive to sample all species (or other taxonomic groupings) of interest within the target communities, with a minimum of nine trait observations per species and study unit (e.g. site, experimental treatment).
4. For each species and trait sampling unit, collect samples randomly across individuals and organs (e.g. leaves) to represent realized community trait distributions.
5. Start by using a nonparametric bootstrapping approach to calculate the moments of the community trait distribution, explicitly incorporating your data structure, experimental design and research question needs into the bootstrapping design.
6. Depending on the moment of interest and sampling design, consider also reporting moment estimates and associated confidence intervals from the other methods. These methods may also better match your system/research design assumptions.
7. Do not only assess shifts in moments independently. Each moment does not represent independent aspects of an assemblage's trait variation or functional variation. Non-independence generates a mathematical constraint on the possible value moments observed for a community (Gross et al., 2017). Trait diversity can thus be assessed with multiple moments. For example, as the trait-abundance distributions become more skewed, they also become more peaked, indicating a decrease in trait diversity.
8. For multivariate metrics, make sure the estimators used are efficient, unbiased and consistent.
9. Read the literature with scepticism. Our results suggest that studies reporting community-weighted moment estimates may be biased depending on the method and sample sizes. For example, our analyses suggest that reports in the literature of (i) CWV are generally biased to lower values (see Figure S5), whereas (ii) CWS and CWK are biased towards higher values (see Figure S5) and (iii) these biases may persist at all sample sizes (see Figure S11). Likewise, our analyses of multivariate metrics suggest that even the rank order of some common metrics cannot be accurately estimated using either a community-weighted or bootstrapping approach (see Figure S17).

3.6 | Caveats and considerations: Every rose has its thorn

While our results strongly suggest that nonparametric bootstrapping is, in general, the best option for estimating the shapes of trait distributions, there may be specific cases where other options are preferable or at least comparable. Estimates of mean trait values tended to be similar across bootstrapping and community-weighted approaches, and analyses using these approaches will likely yield similar results. However, bootstrapping, and by extension, the `TRAITSTRAP` package, provide the user with added benefits through an estimate of uncertainty (CIs) and automatic trait selection. In some situations, parametric bootstrapping may be preferable, for

example, in cases where we have strong a priori assumptions of distribution shapes or because we are explicitly interested in values beyond those observed. However, while we have strong reasons to assume that body size data follow a lognormal distribution (Kerkhoff & Enquist, 2009), for example, our analyses reveal that nonparametric bootstrapping still performs better and would still be the better option for a majority of the size-related traits across our example datasets (Figures S6–S8).

We note that the performance of parametric bootstrapping will also depend on the approach used to fit the distributions and the correspondence between the parametric distribution fitted and the true distribution. Like bootstrapping, Bayesian statistical approaches offer solutions for incorporating intraspecific variation and accounting for uncertainty in estimates (Funk et al., 2017; Shiklomanov et al., 2020), and are conceptually related to bootstrapping (Efron, 2011; Newton & Raftery, 1994; Rubin, 1981). While here we focus on bootstrapping due to its ease of implementation and computational speed, there may be situations where Bayesian approaches are preferable or even necessary. In particular, hierarchical Bayesian models may be well suited for characterizing community trait distributions given the hierarchical nature of community trait distributions (i.e. the community trait distribution is the sum of one or more species trait distributions; Webb et al., 2010). Furthermore, hierarchical Bayesian models can easily incorporate abiotic or biotic covariates that may impact trait distributions, handle missing data, incorporate measurement error and show better performance with rare species (Dietze et al., 2008). Future work using simulations to examine the performance of hierarchical Bayesian models across different sampling scenarios, particularly in comparison to bootstrapping and community-weighted approaches, is needed.

A useful estimator should be consistent, efficient and unbiased (Everitt & Skrondal, 2010). Bootstrapping methods outperformed community-weighted approaches based on these criteria, with community-weighted methods often not consistent, instead even diverging from the true value as sample size increased (Figure 6, Figures S7, S8, S12 and S14) and often showing strong biases (Figures S3 and S4). These poor performances are due to the aggregation biases inherent in community weighting. Community-weighted estimates diverge from the truth because, at low sample sizes, they may, by chance, get closer to the truth, but at larger sample sizes, they converge on the wrong answer. These approaches are inherently biased because they ignore an important source of variation. For some of the multidimensional metrics, neither bootstrapping nor community-weighted approaches proved a good estimator (e.g. FEve, FRic), suggesting that the use of these metrics should be reconsidered until good estimators are developed (Everitt & Skrondal, 2010).

Finally, we stress that there may be valid reasons to focus on the average traits of a subset of individuals or organs within a species, such as providing a more standardized point of comparison or in cases where the research question focuses on potential rather than realized trait expressions. Such analyses ask different questions than those focusing on full trait distributions, and what is appropriate data for the former will lead to bias in the latter.

4 | CONCLUSIONS

Central to several core questions in evolutionary biology, ecology and biogeography are how ecological and evolutionary processes both influence and are influenced by the shape and dynamics of phenotypic distributions (Garnier et al., 2016; Lynch & Walsh, 1998; Schimper, 1903; Walsh & Lynch, 2018). Using community weighting has been a common practice in trait-based ecology to estimate phenotypic distributions' moments (and, therefore, shapes). A problem with this approach is that community weighting leads to an aggregation bias termed the ecological fallacy (Clark & Avery, 2010; Denny, 2017; Ruel & Ayres, 1999). Community weighting ignores important characteristics of trait distributions—intraspecific variation and uncertainty—and our results underscore that this approach is problematic and likely has led to erroneous results and biased conclusions.

We found that a general bootstrapping approach can better estimate trait (phenotypic) distributions and thus provides a general basis for improving the study and quantification of the diversity of form and function. We compared the performance of multiple forms of bootstrapping and community-weighted approaches to estimating the moments of community trait distributions by conducting simulations on multiple datasets that include plant and animal, terrestrial and aquatic, community and population and experimental and observational data. Our method incorporates intraspecific variability and uncertainty and, as a result, can more accurately estimate the shape (moments) of trait distributions. It provides a powerful alternative to estimating trait distributions, especially when assessing higher moments (i.e. allowing better estimation of shapes of distributions beyond means) and a way to measure uncertainty (95% confidence intervals) for each moment.

Our work also points to five key recommendations/conclusions on the status of how to best estimate the shape of phenotypic distributions based on current methods. First, it is important to design trait sampling to reliably quantify the distribution of phenotypes *within* populations and communities of interest (i.e. adhere to the 'rule of 9', and sample randomly unless the research question specifically requires otherwise). Second, make sure to understand and be critical of biases inherent in traditional community-weighted approaches. These approaches bias estimates of variance towards lower values (see Figure S5) and estimates of skewness and kurtosis towards higher values (see Figure S6). Third, when sampling traits or compiling trait data, it is critical to sample as locally as possible to the study site and/or from similar ecological conditions. Using traits from global trait compilations or more distant samples will increasingly inject error and possible bias to your study. Studies that were based on global trait data should likely be revisited. Fourth, robust and accurate measures of trait diversity and multifunctionality necessitate measures of both skewness and kurtosis of trait distributions (Gross et al., 2017). As a result, it is important to sample traits locally and include multiple intraspecific trait measures. Studies on functional diversity that do not include intraspecific trait data are likely suspect. Lastly, none of the multivariate metrics

investigated here (FEve, FDis, RaoQ, FRic, FDiv, average functional uniqueness and average functional), or even their relative ordering, could be accurately estimated using a community-weighted approach (Figures S16 and S17). While bootstrapping methods performed better, some metrics could not be accurately estimated using either bootstrapping or community weighting (FEve, FRic, FDiv and average uniqueness; see Figures S16 and S17). Studies of community trait distributions using species mean traits and multidimensional metrics are thus also likely suspect.

Our results underscore the importance of intraspecific variation in addressing core questions in ecology and evolution (Agrawal, 2001; Bolnick et al., 2011; Finney et al., 2002; Hendry, 2016; Miner et al., 2005; Parmesan & Yohe, 2003; Reznick & Ghalambor, 2001; Violle et al., 2012). Past efforts that have assessed how evolutionary, ecological or environmental forces have shaped trait distributions using community-weighted approaches and biased sampling should likely be re-evaluated in light of our findings. Future efforts aimed at assessing how phenotypic traits are shaped by ecological, environmental or historical factors, or how traits impact ecosystem processes, will likely be enhanced by additional attention to the higher moments of trait distributions, improved sampling designs incorporating both intraspecific and interspecific variation and the use of non-parametric bootstrapping.

AUTHOR CONTRIBUTIONS

Brian S. Maitner, Aud H. Halbritter, Richard J. Telford, Vigdis Vandvik and Brian J. Enquist designed the study. Richard J. Telford, Aud H. Halbritter and Brian S. Maitner provided new methods. Brian S. Maitner performed simulations. Tanya Strydom visualized results and designed the conceptual figures. Christine Lamanna, Lindsey L. Sloat, Andrew J. Kerkhoff, Brian J. Enquist, Julie Messier, Nick Rasmussen, Francesco Pomati and Ewa Merz provided the data used in this manuscript. Brian S. Maitner, Aud H. Halbritter, Tanya Strydom, Vigdis Vandvik and Brian J. Enquist wrote the initial draft, and all authors contributed to revisions.

ACKNOWLEDGEMENTS

B.S.M.'s ugly and inefficient R code that left the other co-first authors (R.T. and A.H.) no choice but to write the TRAITSTRAP package. Funding was provided by the Norwegian Research Council grant ExperTS, project 287784 under the INTPART programme. B.J.E. and B.S.M. were supported by NSF HDR-1934790. B.J.E. was supported by NSF BoCP-2225076. B.J.E., B.S.M. and J.C. were supported by, and several of our plant trait datasets used in our analyses are based on work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Earth and Environmental Systems Sciences Division and Data Management Program, under Award Number DEAC02-05CH11231, as part of the Watershed Function Scientific Focus Area and the ExaShed Project. The plankton dataset was funded by the Swiss National Science Foundation (project num. 182124). The authors would like to thank Carlos P. Carmona and two anonymous reviewers for their comments which greatly improved the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14160>.

DATA AVAILABILITY STATEMENT

The datasets used in this manuscript are publicly available via Zenodo: <https://doi.org/10.5281/zenodo.7876646> (Maitner et al., 2023) and Github (<https://github.com/weecology/portalr>). All code used in this manuscript is publicly available on Github (https://github.com/EnquistLab/trait_bootstrap_ms). The TRAITSTRAP R package is available via CRAN (<https://cran.r-project.org/package=traitstrap>).

ORCID

Brian S. Maitner  <https://orcid.org/0000-0002-2118-9880>

Aud H. Halbritter  <https://orcid.org/0000-0003-2597-6328>

Richard J. Telford  <https://orcid.org/0000-0001-9826-3076>

Tanya Strydom  <https://orcid.org/0000-0001-6067-1349>

Julia Chacon  <https://orcid.org/0000-0002-6787-1295>

Christine Lamanna  <https://orcid.org/0000-0002-6773-7352>

Lindsey L. Sloat  <https://orcid.org/0000-0002-2986-9725>

Andrew J. Kerkhoff  <https://orcid.org/0000-0001-9617-8687>

Julie Messier  <https://orcid.org/0000-0003-1637-6793>

Nick Rasmussen  <https://orcid.org/0000-0002-5106-9038>

Francesco Pomati  <https://orcid.org/0000-0002-9746-3735>

Ewa Merz  <https://orcid.org/0000-0001-8699-9414>

Vigdis Vandvik  <https://orcid.org/0000-0003-4651-4798>

Brian J. Enquist  <https://orcid.org/0000-0002-6124-7096>

REFERENCES

- Abrams, P. A., & Chen, X. (2002). The effect of competition between prey species on the evolution of their vulnerabilities to a shared predator. *Evolutionary Ecology Research*, 4, 897–909.
- Agrawal, A. A. (2001). Phenotypic plasticity in the interactions and evolution of species. *Science*, 294, 321–326.
- Baraloto, C., Timothy Paine, C. E., Patino, S., Bonal, D., Hérault, B., & Chave, J. (2010). Functional trait variation and sampling strategies in species-rich plant communities. *Functional Ecology*, 24, 208–216.
- Bisi, F., Newey, S., Nodari, M., Wauters, L. A., Harrison, A., Thirgood, S., & Martinoli, A. (2011). The strong and the hungry: Bias in capture methods for mountain hares *Lepus timidus*. *Wildlife Biology*, 17, 311–316.
- Bjorkman, A. D., Myers-Smith, I. H., Elmendorf, S. C., Normand, S., Thomas, H. J. D., Alatalo, J. M., Alexander, H., Anadon-Rosell, A., Angers-Blondin, S., Bai, Y., Baruah, G., te Beest, M., Berner, L., Björk, R. G., Blok, D., Bruelheide, H., Buchwal, A., Buras, A., Carbognani, M., ... Zamin, T. (2018). Tundra Trait Team: A database of plant traits spanning the tundra biome. *Global Ecology and Biogeography*, 27, 1402–1411.
- Blanchet, S., Grenouillet, G., Beauchard, O., Tedesco, P. A., Leprieux, F., Dürr, H. H., Busson, F., Oberdorff, T., & Brosse, S. (2010). Non-native species disrupt the worldwide patterns of freshwater fish

- body size: Implications for Bergmann's rule. *Ecology Letters*, 13, 421–431.
- Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., Rudolf, V. H. W., Schreiber, S. J., Urban, M. C., & Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, 26, 183–192.
- Bolnick, D. I., Svanbäck, R., Fordyce, J. a., Yang, L. H., Davis, J. M., Hulsey, C. D., & Forister, M. L. (2003). The ecology of individuals: Incidence and implications of individual specialization. *The American Naturalist*, 161, 1–28.
- Boonstra, R., & Rodd, F. H. (1982). Another potential bias in the use of the longworth trap. *Journal of Mammalogy*, 63, 672–675.
- Botta-Dukát, Z. (2005). Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *Journal of Vegetation Science*, 16, 533–540.
- Brown, J. H. (1995). *Macroecology*. American Library Association.
- Carmona, C. P., de Bello, F., Mason, N. W. H., & Lepš, J. (2016). Traits without borders: Integrating functional diversity across scales. *Trends in Ecology & Evolution*, 31, 382–394.
- Carmona, C. P., Rota, C., Azcárate, F. M., & Peco, B. (2015). More for less: Sampling strategies of plant functional traits across local environmental gradients. *Functional Ecology*, 29, 579–588.
- Chernick, M. R. (2011). *Bootstrap methods: A guide for practitioners and researchers*. John Wiley & Sons.
- Cianciaruso, M. V., Batalha, M. A., Gaston, K. J., & Petchey, O. L. (2009). Including intraspecific variability in functional diversity. *Ecology*, 90, 81–89.
- Clark, W. A. V., & Avery, K. L. (2010). The effects of data aggregation in statistical analysis. *Geographical Analysis*, 8, 428–438.
- Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press.
- Delalandre, L., Gaüzère, P., Thuiller, W., Cadotte, M., Mouquet, N., Mouillot, D., Munoz, F., Denelle, P., Loiseau, N., Morin, X., & Violle, C. (2022). Functionally distinct tree species support long-term productivity in extreme environments. *Proceedings of the Biological Sciences*, 289, 20211694.
- Denelle, P., Violle, C., & Munoz, F. (2019). Distinguishing the signatures of local environmental filtering and regional trait range limits in the study of trait–environment relationships. *Oikos*, 128, 960–971.
- Denny, M. (2017). The fallacy of the average: On the ubiquity, utility and continuing novelty of Jensen's inequality. *The Journal of Experimental Biology*, 220, 139–146.
- Dietze, M. C., Wolosin, M. S., & Clark, J. S. (2008). Capturing diversity and interspecific variability in allometries: A hierarchical approach. *Forest Ecology and Management*, 256, 1939–1948.
- Donoghue, M. J., & Edwards, E. J. (2014). Biome shifts and niche evolution in plants.
- Duffy, J. E. (2003). Biodiversity loss, trophic skew and ecosystem functioning. *Ecology Letters*, 6, 680–687.
- Echeverría-Londoño, S., Enquist, B. J., Neves, D. M., Violle, C., Boyle, B., Kraft, N. J. B., Maitner, B. S., McGill, B., Peet, R. K., Sandel, B., Smith, S. A., Svenning, J. C., Wiser, S. K., & Kerkhoff, A. J. (2018). Plant functional diversity and the biogeography of biomes in North and South America. *Frontiers in Ecology and Evolution*, 6, 219.
- Efron, B. (1979). *Bootstrap methods: Another look at the jackknife*. The Annals of Statistics.
- Efron, B. (2011). The bootstrap and Markov-chain Monte Carlo. *Journal of Biopharmaceutical Statistics*, 21, 1052–1062.
- Enquist, B. J., Bentley, L. P., Shenkin, A., Maitner, B., Savage, V., Michaletz, S., Blonder, B., Buzzard, V., Espinoza, T. E. B., Farfan-Rios, W., Doughty, C. E., Goldsmith, G. R., Martin, R. E., Salinas, N., Silman, M., Díaz, S., Asner, G. P., Malhi, Y., & Hampe, A. (2017). Assessing trait-based scaling theory in tropical forests spanning a broad temperature gradient. *Global Ecology and Biogeography*, 26, 1357–1373.
- Enquist, B. J., Feng, X., Boyle, B., Maitner, B., Newman, E. A., Jørgensen, P. M., Roehrdanz, P. R., Thiers, B. M., Burger, J. R., Corlett, R. T., Couvreur, T. L. P., Dauby, G., Donoghue, J. C., Foden, W., Lovett, J. C., Marquet, P. A., Merow, C., Midgley, G., Morueta-Holme, N., ... McGill, B. J. (2019). The commonness of rarity: Global and future distribution of rarity across land plants. *Science Advances*, 5, eaaz0414.
- Enquist, B. J., Norberg, J., Bonser, S. P., & Violle, C. (2015). Scaling from traits to ecosystems: Developing a general trait driver theory via integrating trait-based and metabolic scaling theories. *Advances in Ecological*, 52, 249–318.
- Everitt, B. S., & Skrondal, A. (2010). *The Cambridge dictionary of statistics*.
- Felsenstein, J. (1988). Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics*, 19, 445–471.
- Finney, B. P., Gregory-Eaves, I., Douglas, M. S. V., & Smol, J. P. (2002). Fisheries productivity in the northeastern Pacific Ocean over the past 2,200 years. *Nature*, 416, 729–733.
- Fontana, S., Thomas, M. K., Moldoveanu, M., Spaak, P., & Pomati, F. (2018). Individual-level trait diversity predicts phytoplankton community properties better than species richness or evenness. *The ISME Journal*, 12, 356–366.
- Freckleton, R. P., & Watkinson, A. R. (2001). Asymmetric competition between plant species. *Functional Ecology*, 15, 615–623.
- Funk, J. L., Larson, J. E., Ames, G. M., Butterfield, B. J., Cavender-Bares, J., Firn, J., Laughlin, D. C., Sutton-Grier, A. E., Williams, L., & Wright, J. (2017). Revisiting the Holy Grail: Using plant functional traits to understand ecological processes. *Biological Reviews*, 92, 1156–1173.
- Garnier, E., Navas, M.-L., & Grigulis, K. (2016). *Plant functional diversity: Organism traits, community structure, and ecosystem properties*. Oxford University Press.
- Gaüzère, P., Doucier, G., Devicor, V., & Kéfi, S. (2019). A framework for estimating species-specific contributions to community indicators. *Ecological Indicators*, 99, 74–82.
- Grenié, M., Denelle, P., Tucker, C. M., Munoz, F., & Violle, C. (2017). funrar: An R package to characterize functional rarity. *Diversity and Distributions*, 23, 1365–1371.
- Grime, J. P. (1998). Benefits of plant diversity to ecosystems: Immediate, filter and founder effects. *Journal of Ecology*, 86, 902–910.
- Grime, J. P. (2006). Trait convergence and trait divergence in herbaceous plant communities: Mechanisms and consequences. *Journal of Vegetation Science*, 17, 255–260.
- Gross, N., Le Bagousse-Pinguet, Y., Liancourt, P., Berdugo, M., Gotelli, N. J., & Maestre, F. T. (2017). Functional trait diversity maximizes ecosystem multifunctionality. *Nature Ecology & Evolution*, 1, 0132.
- Hart, S. P., Schreiber, S. J., & Levine, J. M. (2016). How variation between individuals affects species coexistence. *Ecology Letters*, 19, 825–838.
- Hendry, A. P. (2016). Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. *The Journal of Heredity*, 107, 25–41.
- Henn, J. J., Buzzard, V., Enquist, B. J., Halbritter, A. H., Klanderud, K., Maitner, B. S., Michaletz, S. T., Pötsch, C., Seltzer, L., Telford, R. J., Yang, Y., Zhang, L., & Vandvik, V. (2018). Intraspecific trait variation and phenotypic plasticity mediate alpine plant species response to climate change. *Frontiers in Plant Science*, 9, 1548.
- Inouye, B. D., Ehrlén, J., & Underwood, N. (2019). Phenology as a process rather than an event: From individual reaction norms to community metrics. *Ecological Monographs*, 89, e01352.
- Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., Sechrest, W., Boakes, E. H., Carbone, C., Connolly, C., Cutts, M. J., Foster, J. K., Grenyer, R., Habib, M., Plaster, C. A., Price, S. A., Rigby, E. A., Rist, J., ... Purvis, A. (2009). PanTHERIA: A species-level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090-184. *Ecology*, 90, 2648.
- Jónsdóttir, I. S., Halbritter, A. H., Christiansen, C. T., Althuizen, I. H. J., Haugum, S. V., Henn, J. J., Björnsdóttir, K., Maitner, B. S., Malhi, Y., Michaletz, S. T., Roos, R. E., Klanderud, K., Lee, H., Enquist, B.

- J., & Vandvik, V. (2022). Intraspecific trait variability is a key feature underlying high Arctic plant community resistance to climate warming. *Ecological Monographs*, *93*, e1555.
- Keddy, P. A. (1992). Assembly and response rules: Two goals for predictive community ecology. *Journal of Vegetation Science*, *3*, 157–164.
- Kerkhoff, A. J., & Enquist, B. (2009). Multiplicative by nature: Why logarithmic transformation is necessary in allometry. *Journal of Theoretical Biology*, *257*, 519–521.
- Kot, M., Lewis, M. A., & van den Driessche, P. (1996). Dispersal data and the spread of invading organisms. *Ecology*, *77*, 2027–2042.
- Lake, J. K., & Ostling, A. (2009). Comment on “Functional traits and niche-based tree community assembly in an Amazonian forest”. *Science*, *324*, 1015; author reply 1015.
- Laliberté, E., & Legendre, P. (2010). A distance-based framework for measuring functional diversity from multiple traits. *Ecology*, *91*, 299–305.
- Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution*, *30*, 314–334.
- Laughlin, D. C., & Messier, J. (2015). Fitness of multidimensional phenotypes in dynamic adaptive landscapes. *Trends in Ecology & Evolution*, *30*, 487–496.
- Lavorel, S., Grigulis, K., McIntyre, S., Williams, N. S. G., Garden, D., Dorrrough, J., Berman, S., Quétiér, F., Thébault, A., & Bonis, A. (2007). Assessing functional diversity in the field—methodology matters! *Functional Ecology*, *22*, 134–147.
- Le Bagousse-Pinguet, Y., Gross, N., Saiz, H., Maestre, F. T., Ruiz, S., Dacal, M., Asensio, S., Ochoa, V., Gozalo, B., Cornelissen, J. H. C., Deschamps, L., García, C., Maire, V., Milla, R., Salinas, N., Wang, J., Singh, B. K., & García-Palacios, P. (2021). Functional rarity and evenness are key facets of biodiversity to boost multifunctionality. *Proceedings of the National Academy of Sciences of the United States of America*, *118*, e2019355118.
- Levins, R. (1968). *Evolution in changing environments: Some theoretical explorations*. Princeton University Press.
- Lynch, M., & Walsh, B. (1998). *Genetics and analysis of quantitative traits*. Sinauer Sunderland.
- Maitner, B. S., Boyle, B., Casler, N., Condit, R., Donoghue, J., Durán, S. M., Guaderrama, D., Hinchliff, C. E., Jørgensen, P. M., Kraft, N. J. B., McGill, B., Merow, C., Morueta-Holme, N., Peet, R. K., Sandel, B., Schildhauer, M., Smith, S. A., Svenning, J.-C., Thiers, B., ... Enquist, B. J. (2017). The BIEN R package: A tool to access the Botanical Information and Ecology Network (BIEN) database. *Methods in Ecology and Evolution*, *9*, 373–379.
- Maitner, B. S., Halbritter, A. H., Telford, R. J., Strydom, T., Chacon, J., Lamanna, C., Sloat, L. L., Kerkhoff, A. J., Messier, J., Rasmussen, N., Pomati, F., Merz, E., Vandvik, V., & Enquist, B. J. (2023). Data for “Bootstrapping outperforms community-weighted approaches for estimating the shapes of phenotypic distributions.”. *Zenodo*. <https://doi.org/10.5281/zenodo.7876646>
- Mammola, S., Carmona, C. P., Guillerme, T., & Cardoso, P. (2021). Concepts and applications in functional diversity. *Functional Ecology*, *35*, 1869–1885.
- Mason, N. W. H., Lanoiselée, C., Mouillot, D., Wilson, J. B., & Argillier, C. (2008). Does niche overlap control relative abundance in French lacustrine fish communities? A new method incorporating functional traits. *The Journal of Animal Ecology*, *77*, 661–669.
- Mason, N. W. H., Mouillot, D., Lee, W. G., & Wilson, J. B. (2005). Functional richness, functional evenness and functional divergence: The primary components of functional diversity. *Oikos*, *111*, 112–118.
- Maurer, B. A. (1999). *Untangling ecological complexity: The macroscopic perspective*. University of Chicago Press.
- Miller, J. E. D., Damschen, E. I., & Ives, A. R. (2018). Functional traits and community composition: A comparison among community-weighted means, weighted correlations, and multilevel models. *Methods in Ecology and Evolution*, *10*, 415–425.
- Miner, B. G., Sultan, S. E., Morgan, S. G., Padilla, D. K., & Relyea, R. A. (2005). Ecological consequences of phenotypic plasticity. *Trends in Ecology & Evolution*, *20*, 685–692.
- Moncrieff, G. R., Bond, W. J., & Higgins, S. I. (2016). Revising the biome concept for understanding and predicting global change impacts. *Journal of Biogeography*, *43*, 863–873.
- Mouillot, D., Stubbs, W., Faure, M., Dumay, O., Tomasini, J. A., Wilson, J. B., & Chi, T. D. (2005). Niche overlap estimates based on quantitative functional traits: A new family of non-parametric indices. *Oecologia*, *145*, 345–353.
- Münkemüller, T., Gallien, L., Pollock, L. J., Barros, C., Carboni, M., Chalmandrier, L., Mazel, F., Mokany, K., Roquet, C., Smyčka, J., Talluto, M. V., & Thuiller, W. (2020). Dos and don'ts when inferring assembly rules from diversity patterns. *Global Ecology and Biogeography*, *29*, 1212–1229.
- Muscarella, R., & Uriarte, M. (2016). Do community-weighted mean functional traits reflect optimal strategies? *Proceedings of the Biological Sciences*, *283*, 20152434.
- Navas, M., & Violle, C. (2009). Plant traits related to competition: How do they shape the functional diversity of communities? *Community Ecology*, *10*, 131–137.
- Newton, M. A., & Raftery, A. E. (1994). Approximate Bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society*, *56*, 3–26.
- Norberg, J., Swaney, D. P., Dushoff, J., Lin, J., Casagrandi, R., & Levin, S. A. (2001). Phenotypic diversity and ecosystem functioning in changing environments: A theoretical framework. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 11376–11381.
- Oliveira, B. F., São-Pedro, V. A., Santos-Barrera, G., Penone, C., & Costa, G. C. (2017). AmphioBIO, a global database for amphibian ecological traits. *Scientific Data*, *4*, 170123.
- Paine, C. E. T., Baraloto, C., & Díaz, S. (2015). Optimal strategies for sampling functional traits in species-rich forests. *Functional Ecology*, *29*, 1325–1331.
- Parnesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*, 37–42.
- Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H., & Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: Which approach performs the best? *Methods in Ecology and Evolution*, *5*, 961–970.
- Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte, M. S., Cornwell, W. K., Craine, J. M., Gurvich, D. E., Urcelay, C., Veneklaas, E. J., Reich, P. B., Poorter, L., Wright, I. J., Ray, P., Enrico, L., Pausas, J. G., de Vos, A. C., ... Cornelissen, J. H. C. (2013). New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany*, *61*, 167–234.
- Petchey, O. L., & Gaston, K. J. (2002). Functional diversity (FD), species richness and community composition. *Ecology Letters*, *5*, 402–411.
- Pistón, N., Bello, F., Dias, A. T. C., Götzenberger, L., Rosado, B. H. P., Mattos, E. A., Salguero-Gómez, R., & Carmona, C. P. (2019). Multidimensional ecological analyses demonstrate how interactions between functional traits shape fitness and life history strategies. *Journal of Ecology*, *107*, 2317–2328.
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Reznick, D. N., & Ghalambor, C. K. (2001). The population ecology of contemporary adaptations: What empirical studies reveal about the conditions that promote adaptive evolution. *Genetica*, *112–113*, 183–198.
- Rosenzweig, M. L., & Lomolino, M. V. (1997). Who gets the short bits of the broken stick? In W. E. Kunin & K. J. Gaston (Eds.), *The biology of rarity: Causes and consequences of rare–common differences* (pp. 63–90). Springer.

- Ross, S. R. P.-J., Hassall, C., Hoppitt, W. J. E., Edwards, F. A., Edwards, D. P., & Hamer, K. C. (2017). Incorporating intraspecific trait variation into functional diversity: Impacts of selective logging on birds in Borneo. *Methods in Ecology and Evolution*, 8, 1499–1505.
- Roughgarden, J. (1972). Evolution of niche width. *The American Naturalist*, 106, 683–718.
- Roughgarden, J. (1974). Niche width: Biogeographic patterns among anolis lizard populations. *The American Naturalist*, 108, 429–442.
- Roughgarden, J. (1979). *Theory of population genetics and evolutionary ecology: An introduction*. Macmillan.
- Rubin, D. B. (1981). The Bayesian bootstrap. *The Annals of Statistics*, 9, 130–134.
- Ruel, J. J., & Ayres, M. P. (1999). Jensen's inequality predicts effects of environmental variation. *Trends in Ecology & Evolution*, 14, 361–366.
- Schimper, A. F. W. (1903). Plant geography upon a physiological basis.
- Shackell, N. L., Frank, K. T., Fisher, J. A. D., Petrie, B., & Leggett, W. C. (2010). Decline in top predator body size and changing climate alter trophic structure in an oceanic ecosystem. *Proceedings of the Royal Society B: Biological Sciences*, 277, 1353–1360.
- Shiklomanov, A. N., Cowdery, E. M., Bahn, M., Byun, C., Jansen, S., Kramer, K., Minden, V., Niinemets, Ü., Onoda, Y., Soudzilovskaia, N. A., & Dietze, M. C. (2020). Does the leaf economic spectrum hold within plant functional types? A Bayesian multivariate trait meta-analysis. *Ecological Applications*, 30, e02064.
- Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., Aarssen, L. W., Baraloto, C., Carlucci, M. B., Cianciaruso, M. V., L. Dantas, V., Bello, F., Duarte, L. D. S., Fonseca, C. R., Freschet, G. T., Gaucherand, S., Gross, N., Hikosaka, K., Jackson, B., ... Wardle, D. A. (2015). A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters*, 18, 1406–1419.
- Spehn, E. M., Hector, A., Joshi, J., Scherer-Lorenzen, M., Schmid, B., Bazeley-White, E., Beierkuhnlein, C., Caldeira, M. C., Diemer, M., Dimitrakopoulos, P. G., Finn, J. A., Freitas, H., Giller, P. S., Good, J., Harris, R., Höglberg, P., Huss-Danell, K., Jumpponen, A., Koricheva, J., ... Lawton, J. H. (2005). Ecosystem effects of biodiversity manipulations in European grasslands. *Ecological Monographs*, 75, 37–63.
- Suding, K. N., Lavorel, S., Chapin, F. S., Cornelissen, J. H. C., Díaz, S., Garnier, E., Goldberg, D., Hooper, D. U., Jackson, S. T., & Navas, M.-L. (2008). Scaling environmental change through the community-level: A trait-based response-and-effect framework for plants. *Global Change Biology*, 14, 1125–1140.
- Swenson, N. G. (2014). Phylogenetic imputation of plant functional trait databases. *Ecography*, 37, 105–110.
- Todd, B. D., Scott, D. E., Pechmann, J. H. K., & Whitfield Gibbons, J. (2011). Climate change correlates with rapid delays and advancements in reproductive timing in an amphibian community. *Proceedings of the Royal Society B: Biological Sciences*, 278, 2191–2197.
- Turelli, M., & Barton, N. H. (1994). Genetic and statistical analyses of strong selection on polygenic traits: What, me normal? *Genetics*, 138, 913–941.
- Uriarte, M., & Menge, D. (2018). Variation between individuals fosters regional species coexistence. *Ecology Letters*, 21, 1496–1504.
- Villéger, S., Mason, N. W. H., & Mouillot, D. (2008). New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89, 2290–2301.
- Violle, C., Enquist, B. J., McGill, B. J., Jiang, L., Albert, C. H., Hulshof, C., Jung, V., & Messier, J. (2012). The return of the variance: Intraspecific variability in community ecology. *Trends in Ecology & Evolution*, 27, 244–252.
- Violle, C., Reich, P. B., Pacala, S. W., Enquist, B. J., & Kattge, J. (2014). The emergence and promise of functional biogeography. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 13690–13696.
- Violle, C., Thuiller, W., Mouquet, N., Munoz, F., Kraft, N. J. B., Cadotte, M. W., Livingstone, S. W., & Mouillot, D. (2017). Functional rarity: The ecology of outliers. *Trends in Ecology & Evolution*, 32, 356–367.
- Walsh, B., & Lynch, M. (2018). *Evolution and selection of quantitative traits*. Oxford University Press.
- Webb, C. T., Hoeting, J. A., Ames, G. M., Pyne, M. I., & LeRoy Poff, N. (2010). A structured and dynamic framework to advance traits-based theory and prediction in ecology. *Ecology Letters*, 13, 267–283.
- Weber, M. G., & Agrawal, A. A. (2012). Phylogeny, ecology, and the coupling of comparative and experimental approaches. *Trends in Ecology & Evolution*, 27, 394–403.
- Weiher, E., & Keddy, P. (1999). Assembly rules as general constraints on community composition. In *Ecological assembly rules: Perspectives, advances, retreats* (pp. 251–271). Cambridge University Press.
- Weiher, E., & Keddy, P. A. (1995). Assembly rules, null models, and trait dispersion: New questions from old patterns. *Oikos*, 74, 159–164.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4, 1686.
- Wieczynski, D. J., Boyle, B., Buzzard, V., Duran, S. M., Henderson, A. N., Hulshof, C. M., Kerckhoff, A. J., McCarthy, M. C., Michaletz, S. T., Swenson, N. G., Asner, G. P., Bentley, L. P., Enquist, B. J., & Savage, V. M. (2019). Climate shapes and shifts functional biodiversity in forests worldwide. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 587–592.
- Wissinger, S. A. (1992). Niche overlap and the potential for competition and intraguild predation between size-structured populations. *Ecology*, 73, 1431–1444.
- Wong, M. K. L., & Carmona, C. P. (2021). Including intraspecific trait variability to avoid distortion of functional diversity and ecological inference: Lessons from natural assemblages. *Methods Ecology and Evolution*, 12, 946–957.
- Woodward, F. I., Lomas, M. R., & Kelly, C. K. (2004). Global climate and the distribution of plant biomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359, 1465–1476.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix 1. Interpreting community trait distributions.

Appendix 2. Calculating community-weighted moments.

Appendix 3. Bootstrapping approaches.

Appendix 4. Traitstrap package overview.

Appendix 5. Traitstrap vignette.

Appendix 6. Supplementary analyses and results.

Appendix 7. Datasets.

Appendix 8. Simulation methods.

Appendix 9. Supplementary discussion.

Appendix 10. Global plant trait data references.

How to cite this article: Maitner, B. S., Halbritter, A. H., Telford, R. J., Strydom, T., Chacon, J., Lamanna, C., Sloat, L. L., Kerckhoff, A. J., Messier, J., Rasmussen, N., Pomati, F., Merz, E., Vandvik, V., & Enquist, B. J. (2023). Bootstrapping outperforms community-weighted approaches for estimating the shapes of phenotypic distributions. *Methods in Ecology and Evolution*, 00, 1–19. <https://doi.org/10.1111/2041-210X.14160>