This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Data Paper published in *Ecology*, an Ecological Society of America journal

Metadata S1

GLOSSAQUA: A global dataset of size spectra across aquatic ecosystems

Authors: Zeynep Ersoy, Charlotte Evangelista, Aitor Larrañaga, Daniel M. Perkins, Javier Sánchez-Hernández, Teofana Chonova, David Cunillera-Montcusí, Carmen García-Comas, Jorge García-Girón, Ioar de Guzman, Justin Pomeranz, Victor Saito, Matías Arim, Dirceu Baumgartner, Gilmar Baumgartner, Mauro Berazategui, Dani Boix, Giovanna Collyer, Jordi Compte, Almir Manoel Cunico, Renee M. van Dorst, Jon Harding, Ursula Gaedke, Stéphanie Gascón, Éder André Gubiani, Daniel Hernández, James R. Junker, Mercedes López-Vázquez, Anderson Luís Maciel, Thomas Mehner, Roger Paulo Mormul, Ramiro Pereira-Garbero, Danielle Petsch, Pitágoras Augusto Piana, Xavier D. Quintana, Julia Reiss, Lucía Rodríguez-Tricot, Jordi Sala, Wilson Sebastián Serra, Tadeu Siqueira, Helen J. Warburton, Matías Zarucki, Ignasi Arranz

Introduction

The current biodiversity crisis has led to global efforts aimed at collecting new information or compiling existing databases on species richness and abundance across spatial and temporal scales (Magurran et al. 2010; Dornelas et al. 2018; Comte et al. 2021). There is now a growing interest in using trait-based approaches to better predict changes in functional biodiversity, community structure and ecosystem multifunctionality (Pawar et al. 2015; Gibert et al. 2015; Keddy and Laughlin 2021). Among the many measurable biological traits, body size is particularly relevant because it (i) is relatively easy to measure, (ii) correlates with many vital functional features such as metabolism, growth, survival and reproduction (Peters 1983; Sprules and Munawar 1986; Brown et al. 2004), and (iii) has been shown to often respond to environmental disturbances (Deutsch et al. 2022; Liu et al. 2023; Atkinson et al. 2024). In addition, changes in body size can modulate population dynamics and trophic interactions via demographic and predator mouth gape constraints, affecting ecosystem structure and functions such as energy transfer and decomposition rates (Gaedke, 1992; Woodward et al. 2005; García–Comas et al. 2016; Mehner et al. 2022; Larrañaga et al. 2023). Although body size has emerged as a key ecological indicator for studying the effects of global change (Petchey and Belgrano 2010; Basset et al. 2012; Marin et al. 2023), a standardized database of body size structure spanning various ecosystems and taxonomic groups is still missing. This is a necessary step towards developing a finer mechanistic understanding of how environmental and anthropogenic changes affect community size structure.

The size spectrum is the relationship between abundance (or biomass) and individual body size, often on a log-log scale, and independent of species composition (Ghilarov 1944; Sheldon et al. 1972). Traditionally, the size spectrum is binned into arbitrarily defined body size classes in a geometric series (Platt and Denman 1977; Silvert and Platt 1978); however, the parameters estimated may be sensitive to the bin width selected (Edwards et al. 2017) and to the size metric considered (e.g., biomass, biovolume or length) (Sprules and Barth 2016). In addition, size spectra fitted with binning-based methods are usually normalized by dividing the abundance (or biomass) by the width of the size classes (Sprules and Barth 2016). To avoid the bias introduced by the use of binned data, more recent approaches have shown that fitting the size spectrum with either a Pareto probability distribution or estimating parameters by maximum likelihood estimation (MLE) can provide more robust and accurate estimates of exponents than binning-based methods (Vidondo et al. 1997; White et al. 2007; Arim et al. 2011; Edwards et al. 2020). Overall, the size spectrum is defined by three parameters: the slope (which is related to the exponent; Sprules and Barth 2016; Edwards et al. 2017), the intercept and the linearity (defined by the R2 of the linear fit of the size spectrum; see details below). These parameters have been used to interpret patterns of energy fluxes and total biomasses in terrestrial (Potapov et al. 2021) and aquatic (Gaedke 1993; Gaedke and Straile 1994; Kerr and Dickie 2001) food webs. Specifically: (i) the slope of the size spectrum represents the relative proportion of small- vs. large-bodied individuals, with steeper slopes (more negative values) reflecting communities dominated by relatively more small-bodied organisms (Sprules and Barth 2016); (ii) the intercept of the size spectrum can inform about the total biomass within the ecosystem (Jennings and Blanchard 2004); and (iii) the extent of non-linearity can provide information on trophic strategy (e.g., omnivory), ecosystem stability (e.g., predator-prey size ratios) and cross-ecosystem interactions (e.g., resource subsidies) (Chang et al. 2014; Arranz et al. 2019; Perkins et al. 2021).

A growing number of studies have recently used size spectrum parameters to evaluate the impacts of anthropogenic disturbances such as biological invasions (Arranz et al. 2021), climate warming (Dossena et al. 2012; Pomeranz et al. 2022), pollution (Peralta-Maraver et al. 2019), harvesting (Jennings and Blanchard 2004) and land use (Collyer et al. 2023; Arranz et al. 2023; Larrañaga et al. 2023) on aquatic communities. However, there have only been a few attempts to compile large datasets on size spectra to draw robust conclusions about the processes regulating the size structure and functioning of ecosystems (dos Santos et al. 2017; Hatton et al. 2021; Kiko et al. 2022). In one of those compilations, using a systematic literature review, dos Santos et al. (2017) showed that size spectra exhibited steeper slopes and higher intercepts in disturbed compared to undisturbed aquatic ecosystems. Although pioneering, this study was limited to 37 dichotomous studies (disturbed vs. undisturbed environments) and therefore did not account for studies conducted along natural environmental gradients. More recently, Hatton et al. (2021) compiled a size spectrum in the ocean from bacteria to whales, to show that the impact of commercial fishing is the main factor affecting marine food webs. Similarly, Kiko et al. (2022) compiled the first global dataset of particle size distributions in marine ecosystems, including all living (plankton communities) and non-living particles (detritus aggregates or fecal pellets) collected with underwater imaging systems. To date, no global dataset of size spectrum parameters spanning multiple aquatic ecosystems and taxonomic groups exists.

Here, we have assembled the most comprehensive collection of size spectrum parameters across different taxonomic groups inhabiting freshwater, marine and brackish ecosystems. The resulting GLOSSAQUA dataset has been created by combining and curating data from peer-reviewed articles and grey literature (i.e., unpublished datasets from academic research, Foo et al. 2021). Notably, the dataset includes the geographical and ecological information of the study sites, as well as methodological information on the sampling technique and fitting method used for the size spectrum models (i.e., number of size classes, binned or MLE approaches). GLOSSAQUA can be used to, for example, (i) compare global size spectrum patterns among different aquatic ecosystems (e.g., marine vs. freshwater, lentic vs. lotic), taking into account different taxonomic groups, (ii) reveal spatial (e.g., along a latitudinal gradient) and temporal (using the time series and/or sampling years variable) changes in community size structure, and (iii) assess changes in the community size structure along natural environmental gradients from impacted towards less impacted/pristine sites. Finally, because we have provided the geographical coordinates of the study sites/areas, GLOSSAQUA can be used in combination with additional data uploaded from existing databases (e.g., temperature, human footprint, elevation and harvesting pressure) to assess the response of community size structure to global changes. Thus, GLOSSAQUA aims to provide a robust, open-source information baseline for macroecological and biogeographical analyses of aquatic community size spectrum at a global scale.

Class I. Data Set Descriptors

A. Data set identity: GLOSSAQUA: A global dataset of size spectra across aquatic ecosystems

B. Data set identification code: Database accession numbers or site-specific codes used to uniquely identify data set

- Data_S1.zip contains following files:
    - data folder; GLOSSAQUA dataset
        - GLOSSAQUA_DataDictionary.txt
        - GLOSSAQUA_DataSource.txt
        - GLOSSAQUA_Sample.txt
        - GLOSSAQUA_Size.txt
    - Metadata S1 (this file)
- The associated Zenodo release (Ersoy, Evangelista and Arranz, 2025; https://doi.org/10.5281/zenodo.14701391) contains the following files:
    - rscripts folder; R code to reproduce summary figures in the Metadata document and standardize size spectrum slopes
        - Figures_GLOSSAQUA_dataset.R
        - Standardization_GLOSSAQUA_dataset.R
    - GLOSSAQUA_dataset.Rproj; RStudio project to enable use of dataset easily
    - README.md; instructions for the use of the data and the code

C. Data set description

1. Originators:

Zeynep Ersoy. Departamento de Biología y Geología, Física y Química Inorgánica, Universidad Rey Juan Carlos (URJC), Móstoles, Spain.

Charlotte Evangelista. Norwegian Institute for Nature Research (NINA), Trondheim, Norway.

Ignasi Arranz. Instituto de Investigación en Cambio Global (IICG-URJC), Universidad Rey Juan Carlos, Tulipán s/n, 28933 Móstoles, España. Departamento de Biología y Geología, Física y Química Inorgánica, Universidad Rey Juan Carlos (URJC), Móstoles, Spain.

2. Abstract: Body size is a key trait in ecology due to its influence on metabolism and many other life-history traits that affect population and community responses to environmental variation as well as ecosystem properties. The size spectrum represents the relationship between abundance (or biomass) and body size, independent of species identity. Size spectrum parameters, such as the slope or intercept, have been applied extensively as indicators of ecological status across multiple ecosystem types. The GLOSSAQUA dataset includes size spectrum data from mainly heterotrophic communities composed of single (e.g., zooplankton, macroinvertebrates or fish) to multiple taxonomic groups (e.g., from primary consumers to apex predators, and phytoplankton to large zooplankton), across diverse spatial and temporal scales, from surveys in freshwater (43% studies), marine (52% studies) and brackish (5% studies) ecosystems. In total, we compiled a unique global dataset of 8,459 size spectrum slopes or exponents, 5,237 intercepts and 4,497 linearity coefficients (i.e., defined by the R2 of the linear fit of the size spectrum) from 127 articles and grey literature (i.e., unpublished datasets). The

current dataset aims to help identify the main drivers shaping aquatic size spectrum parameters at a global scale and contribute to cross-ecosystem comparisons. GLOSSAQUA can serve to explore questions such as factors influencing spatial and temporal dynamics of community size structure, comparing the response of community size structure between natural vs. human-impacted sites, and comparing global patterns in different aquatic ecosystems. We encourage researchers, especially those from underrepresented geographical areas (e.g., South Hemisphere and Asia) to fuel this dataset in the future. The dataset is provided under a CC-BY-NC-S4 4.0 license and users are encouraged to cite this data paper when using the data.

D. Key words/phrases: biodiversity database, body size distribution, community assembly, food web, global scale, multiple surveys.

Class II. Research origin descriptors

A. Overall project description:

1. Identity: Same as in Class I.A.
2. Originators: Same as in Class I.A.
3. Period of study: The data compilation occurred between 2022 and 2024. Data collection spanned the period 1959–2022.
4. Objectives: Data were compiled to assess spatial and temporal gradients of size spectrum parameters of ecological communities and food webs based on a literature review approach.
5. Abstract: Same as in Class I.C.2.

6. Sources of funding: The Iberian Ecological Society (SIBECOL) funded this study through an Advanced Early Career Researchers Project Grant (ACROSS project).

B. Specific subproject description

1. Site description

Site type: It is worth noting here that we are referring to the sites described in the studies included in the GLOSSAQUA dataset. The current dataset encompasses multiple aquatic ecosystems across freshwater (e.g., lakes, ponds, streams and groundwater; in total, 43.4% studies), marine (e.g., coral reef and open ocean; in total, 51.5% studies) and brackish (e.g., lagoons, estuaries and marshes; in total, 5.1% studies) habitats. The main context in which each study was conducted is also reported in the dataset. Specifically, the studies were either conducted in the context of environmental changes, or developed new methods applied to size spectrum analyses. We attributed each study to one or more categories within the context variable based on its screening (Table 1).

Table 1. Definition of the main contexts of the studies from which data were collected or estimated. Note that more than one context can be attributed to the same study.

Natural habitat heterogeneity: when communities are not subject to obvious human impacts. These studies are generally carried out along natural gradients of environmental conditions (e.g., salinity gradients and depth gradients).
Methodology: when the study's context includes comparing methods related to size spectrum analyses or testing a novel method.
Ecosystem productivity: when the study's context is directly linked to the flux of energy or nutrients (e.g., resource subsidies, upwelling process and oceanic current).
Trophic interaction: when the study's context is directly linked to a shift in trophic interaction (e.g., change in top-down and/or bottom-up forces).

Seasonality: when the study's context is directly linked to seasonal or temporal changes (e.g., summer versus winter surveys and diel variation).
Pollution: when the study's context is linked to any form of contamination that enters in the system (e.g., wastewater pollution and eutrophication).
Weather: when the study's context is directly linked to changes in weather conditions (e.g., flooding and water temperature).
Harvesting: when the study's context is linked to human-induced fishing or harvesting of species (e.g., recreational and commercial fisheries).
Introduction: when the study's context is linked to the presence of non-native species.
Landscape modification: when the study's context is linked to any form of habitat modification (e.g., urbanization).

Geography: The current dataset has a global distribution covering most parts of the world, with a latitudinal range of -76.7º to 80.8º (Figure 1A). Specifically, our dataset represents all continents (4.9% studies in Africa, 0.8% studies in Antarctica, 18% studies in Asia; 36.9% studies in Europe, 20.5% studies in North America, 6.6% studies in Oceania, and 12.3% studies in South America) and the major oceans (10% studies in Antarctic Ocean, 10% studies in Arctic Ocean, 30% studies in Atlantic Ocean, 5% studies in Indian Ocean, 10% studies in Mediterranean Sea and 35% studies in Pacific Ocean) within the eight biogeographic realms (as defined by Olson et al. 2001; Figure 1A).

Habitat: Habitat and ecosystem characteristics are reported, as described in the original study. Specifically, habitats refer to marine, freshwater or brackish waters, while ecosystems refer to stream, lake, pond, among others (Table 3).

Geology, landform: This information is not available.

Watersheds, hydrology: This information is not available.

Site history: This information is not available.

Climate: This information is not available.

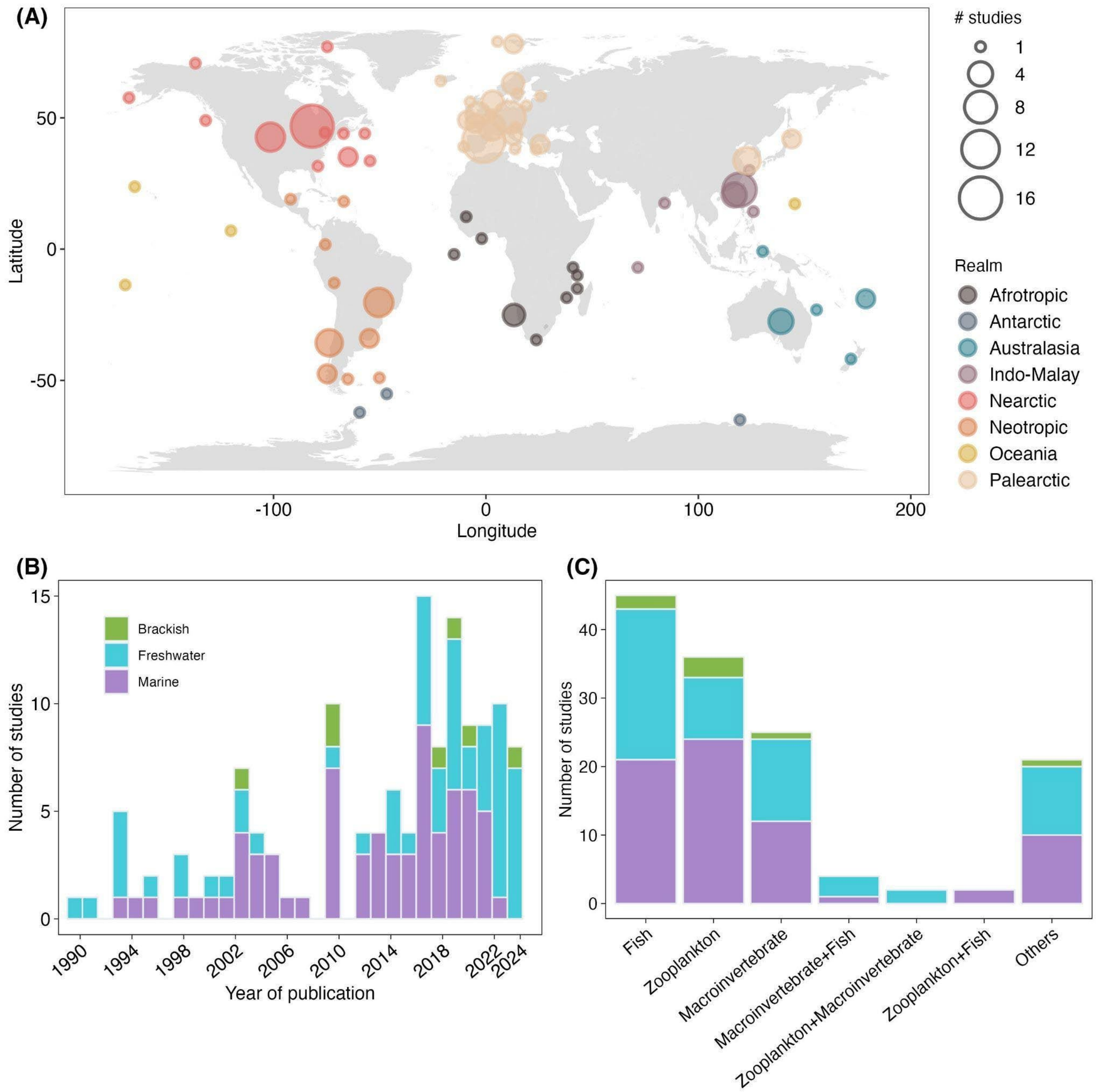

Figure 1. (A) Locations of aquatic studies included in the dataset (n = 135 studies) classified within eight biogeographic realms. The sizes of the circles are proportional to the number of studies per country in each biogeographic realm. Number of studies across (B) years of publication and (C) aquatic taxonomic groups in different aquatic habitats (brackish, freshwater and marine). "Others"

in (C) refers to food webs representing all other combinations of taxonomic groups including primary producers, bacteria and protozoa (e.g., Bacteria + Zooplankton, PrimaryProducer + Protozoa + Macroinvertebrate, Bacteria + PrimaryProducer + Zooplankton and PrimaryProducer + Zooplankton + Fish).

2. Experimental or sampling design

Design characteristics: Data were obtained from two different sources: a systematic review and an online survey (Figure 2). We used the standardized "Preferred Reporting Items for Systematic Reviews and Meta-Analyses" (PRISMA) method to select studies to be included in the dataset (Page et al. 2021). All studies had to meet the following criteria: (i) to be field-based studies and hence, experimental (including in situ experiments) or theoretical studies were excluded because we aim to assess real-world observational patterns rather than seeking for causal links; (ii) to involve animal communities and hence, primary producers were excluded unless they were included as part of the studied food webs (e.g., phytoplankton-zooplankton-fish within the category "Others" in Figure 1C); and (iii) to be studies based on individual body size measurements without relying on mean or maximum body size per species. Importantly, our initial aim was not restricted to aquatic ecosystems a priori (see search string below), but due to the very low number of terrestrial studies (n = 2 after full-text reviews, see details below) we decided to remove them from the dataset.

For the literature review, studies on size spectra were identified through a comprehensive search of Web of Science on 9 March 2022 (Core collection, Environmental Sciences and Ecology topics, Timespan = 1990 and 2022) using the following keywords limited to title and abstract: (size spectr* OR size distribution OR biomass spectr* OR body-size spectr* OR body size spectr* OR size-spectr* OR abundance spectr*) AND (*bacteri* OR *plankto* OR animal* OR *vertebr* OR arthropod* OR amphipod* OR crustace* OR crayfish* OR insect* OR mollus* OR amphibian* OR fish* OR teleost* OR bird* OR avian OR reptile* OR mammal* OR cetace*). We started our search in year 1990 to mitigate potential difficulties in contacting authors of earlier publications and to ensure the quality to digitize figures. Keyword selection was based on the prior size spectrum knowledge of the authors, and it covered all taxonomic groups where size spectrum research has been applied including animal communities (e.g., fish communities) and food webs (e.g., from bacteria to whales).

The Web of Science search identified 8,945 records. After identifying reviews and meta-analyses among these records, we conducted a manual search in these studies' reference lists to retrieve relevant ones that may not have been included in the Web of Science search. This manual search retrieved nine additional records and we excluded three duplicate studies (in total, 8,951 records). Study titles and abstracts were assessed for relevance and we excluded 8,440 records that did not meet our selection criteria (i.e., (i) to be field-based studies; (ii) to involve animal communities and (iii) to be studies based on individual body size measurements without relying on mean or maximum body size per species) or were duplicates (Figure 2). Then, we excluded 384 records via full-text reviews based on the selection criteria previously described, leaving 127 unique studies for the dataset (Figure 2). From these studies, we extracted size spectrum parameters (e.g., slope, intercept and linearity) directly from tables or by digitizing information from plots using WebPlotDigitizer software v.3.4 (Rohatgi 2020). For each study, we also extracted additional information related to geographic location, ecosystem type, ecological context, taxonomic groups, sampling method, sampling year, sampling month – multiple months were considered to build an average size spectrum – and period (sampling year and mainly month within the sampling year), size spectrum method, number of size classes, and minimum and maximum size classes. For the geographical coordinates, we provided site-specific values if available in the articles, but for studies conducted over large areas we provided an average latitude and longitude value. If relevant data or information (e.g., size spectrum parameters and geographic coordinates) were not available for a specific study, we contacted the corresponding authors to request the missing information for inclusion in the dataset.

In addition to the literature review, we launched an online survey through our personal networks, ecological associations, and social media accounts to invite researchers to contribute to expanding GLOSSAQUA by sharing unpublished datasets that provide measurements of body size of individuals within wild communities (Figure 2). We received eight new datasets, all from freshwater ecosystems, for which we calculated the size spectrum parameters following the size spectrum fitting recommendations by Sprules (2022). Specifically, we used dry mass as the measure of body size whenever possible, but if only body length was provided, we transformed body length to mass using length-mass relationships in plankton (Environmental Protection Agency 2010), macroinvertebrates (Benke et al.1999; Méthot et al. 2012) and fish (Tomanova et al. 2010). We then calculated the size spectrum parameters using the MLE approach based on a log-likelihood function (Edwards et al. 2017), but we also computed those parameters using traditional binning techniques (i.e., Normalized Abundance Size Spectrum, Sprules and Barth 2016) to allow for cross-method comparisons.

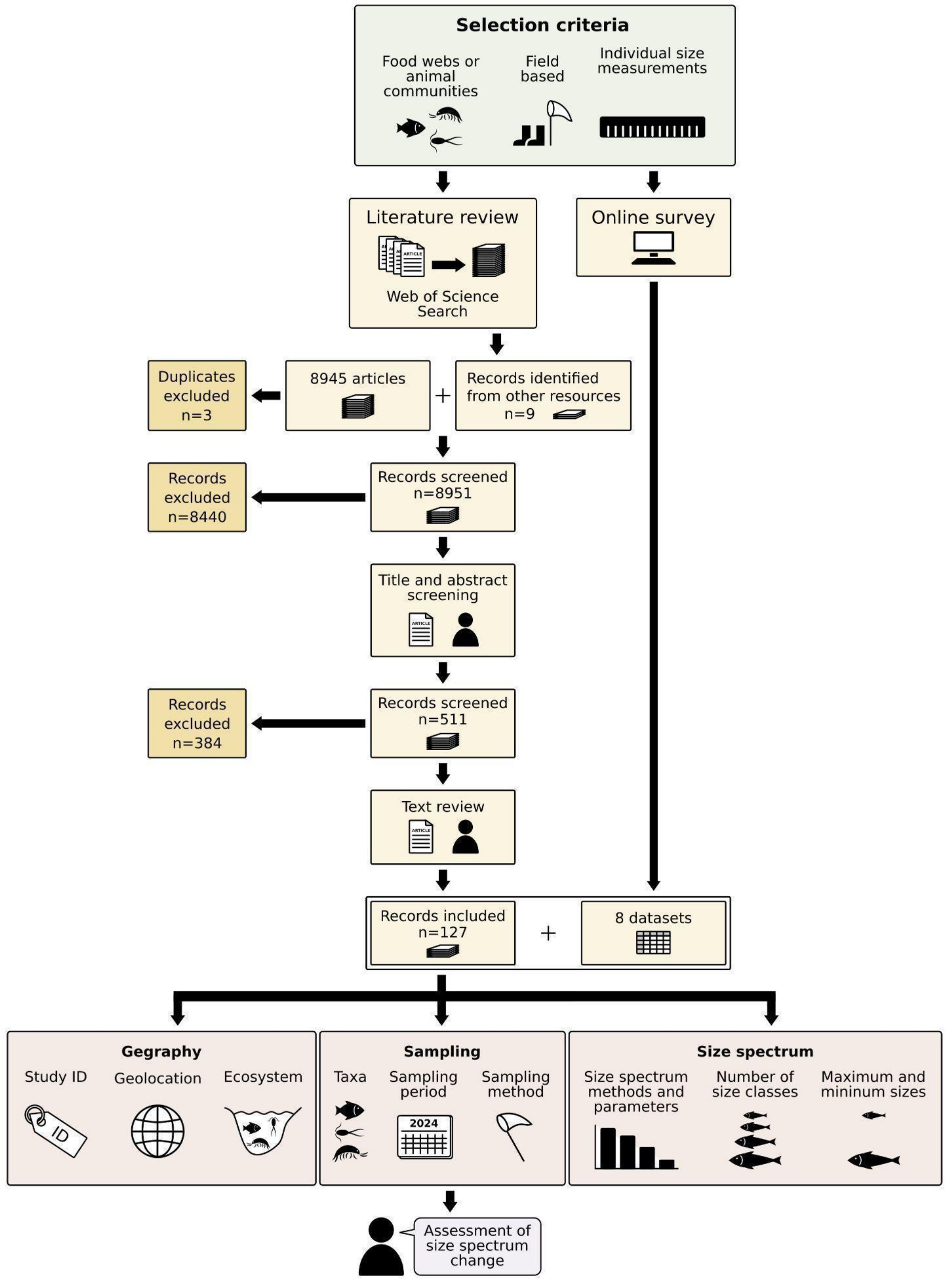

Figure 2. PRISMA diagram showing the results of the search strategy for the GLOSSAQUA dataset. For each step, rectangles on the right show the search records included while rectangles on the left show records that were excluded. We limited our dataset to animal communities with field measurements/observations, using individual body size and with information on the geographical coordinates. Each study included in the dataset is identified by a unique Study ID.

Individual illustrations in the figure were created by Jagoba Malumbres-Olarte and he arranged the components with the support of the principal investigators.

Permanent plots: This information is not available.

Data collection period, frequency, etc.: The interest in the size spectrum research has grown rapidly over the past 30 years (Figure 1B). Data gathered here spanned from 1959 to 2022 and were collected from short- and long-term monitoring, meaning that the frequency of data collection varied among size spectrum studies. In some studies where multiple sampling events occurred, we were able to only extract the size spectrum parameters of the pooled data. GLOSSAQUA covers a large geographical extent and data were collected at different periods, as indicated by the variables "Sampling years" and "Sampling month". The number of times that sites were sampled was also reported in the variable "NumSampled" (see *GLOSSAQUA_Size* file and Table 4 for details), varying across a time range from only 1 sampling to 377 sampling times (e.g., plankton food web in Lake Constance).

3. Research methods

Field/laboratory: Assessing size spectrum variation across aquatic environments is particularly challenging since different ecosystems often use specific survey methods with sampling gears of varying size-selectivity (e.g., electrofishing often has lower catchability of the largest fish than nets)

(Sprules and Barth 2016). For instance, for fish, while pelagic habitats are conventionally sampled using trawls or acoustic techniques, shallow benthic habitats (<2 m) are mainly surveyed using electrofishing, or with trawls and other selective gear. This makes cross-habitat comparisons difficult. By including details on the sampling methodology, the GLOSSAQUA dataset will allow future users to account for its potential influence in the statistical models. Additionally, the values of the size spectrum parameters may vary depending on the method applied to build it and the type of body size measurement (e.g., length, spherical diameter, volume or dry mass). For instance, the size spectrum slope has different theoretical predictions depending on whether abundance or biomass is used (Sprules and Barth 2016). A theoretical value of -2 in Normalized Abundance Size Spectrum (NASS) at a log-log scale is reported when the abundance per size class is used, while the value changes to -1 in Normalized Biomass Size Spectrum (NBSS) when biomass is used (Sprules and Barth 2016). Moreover, different binning techniques (e.g., binning vs. normalized binning or different geometric series) could give different parameter estimates (Edwards et al. 2017). Thus, we annotated the size spectrum methods used under the following six major categories with respect to their binning techniques (i.e., classification of individual body sizes into size classes): (i) Normalized Abundance Size Spectrum (NASS), (ii) Normalized Biomass Size Spectrum (NBSS), (iii) Abundance Size Spectrum (ASS) and (iv) Biomass Size Spectrum (BSS); and the nonbinning techniques: (vi) Maximum Likelihood Estimation (MLE) and (vii) Pareto distributions (i.e., rank/frequency plots) (Figure 3). We also classified the three main types of body size corresponding to body mass, body length and body volume, as suggested by Andersen (2019). In total, we compiled a unique dataset of 8,459 size spectrum slopes or coefficients, 5,237 intercepts and 4,497 linearities (R^2) from 127 articles and eight datasets distributed worldwide (Figure 4). The number of size spectrum parameters differs because of the lack of data in the data source or some size spectrum methods like MLE or Pareto distributions do not allow to directly compute intercepts or R2. Finally, we reported the minimum and maximum body sizes considered to account for size-selective effects and the number of size classes in the case of binning techniques.

Instrumentation: This information is not available in our case.

Taxonomy and systematics: We reported if the studies focused on a set of phylogenetically close species that represent roughly similar body sizes, and generally conform to one feeding type (e.g., fish feeding on zooplankton; hereafter referred to as community); or a set of phylogenetically distant species that show considerable variation in body sizes and feeding types, and where prey-predator dynamics clearly occur within the group (hereafter referred to as food webs). Specifically, at the community level, we identified three main taxonomic groups including zooplankton (i.e., animal organisms drifting with water currents; in total, 33.6% studies), macroinvertebrates (i.e., benthic and pelagic macroinvertebrates; in total, 24.3% studies), and fish (in total, 42.1% studies) (Figure 1C). Additionally, at the food-web level, apart from the combination of the three main taxonomic groups (i.e., zooplankton, macroinvertebrates and fish), there were also different types of food webs with multiple taxonomic groups, such as a food web with bacteria and primary producers or another food web with bacteria, protozoa and macroinvertebrates (labeled as "Others" in Figure 1C).

Permit history: All data from unpublished studies were reviewed to ensure that they had permission from data providers to be included in this current dataset.

Legal/organizational requirements: This information is not available in our case.

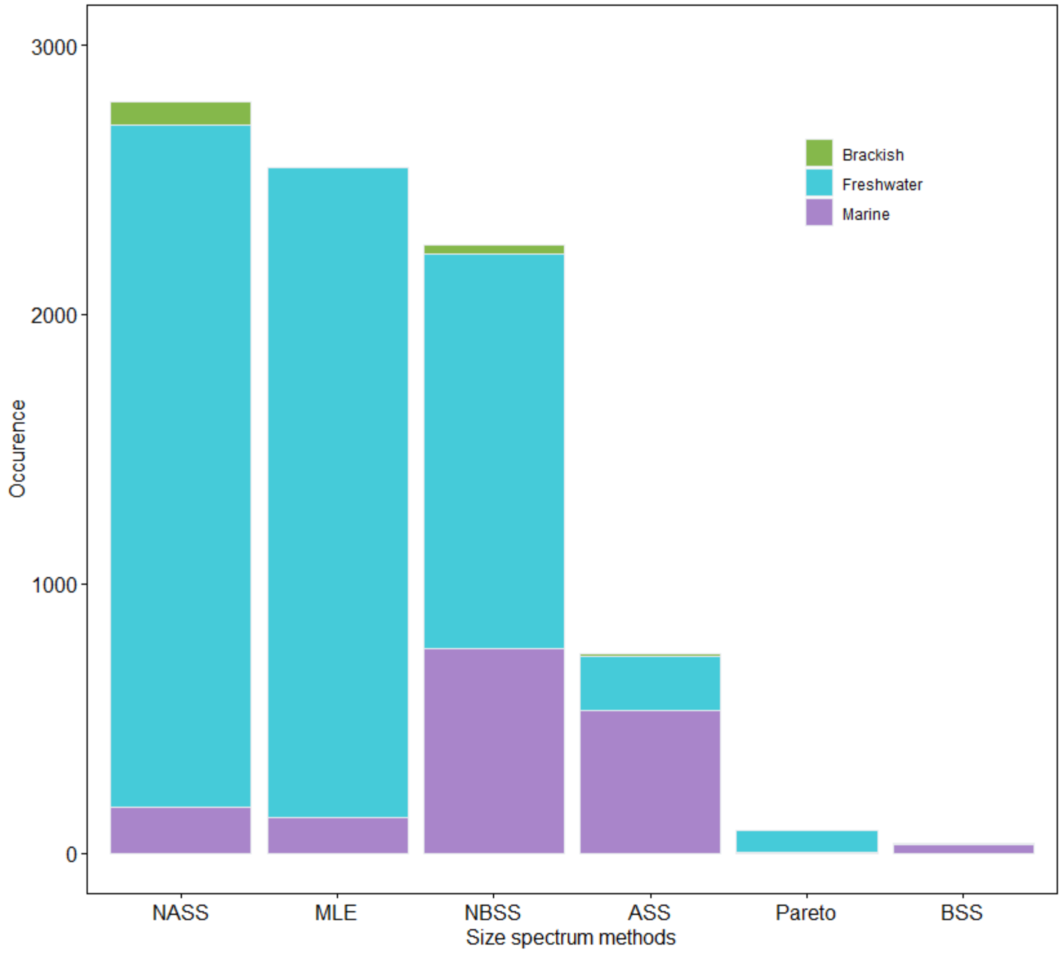

Figure 3. Occurrence of the different size spectrum methods identified in the GLOSSAQUA dataset for aquatic habitats. NASS: Normalized Abundance Size Spectrum; MLE: Maximum Likelihood Estimate; NBSS: Normalized Biomass Size Spectrum; ASS: Abundance Size Spectrum; Pareto: Pareto Distribution; BSS: Biomass Size Spectrum.

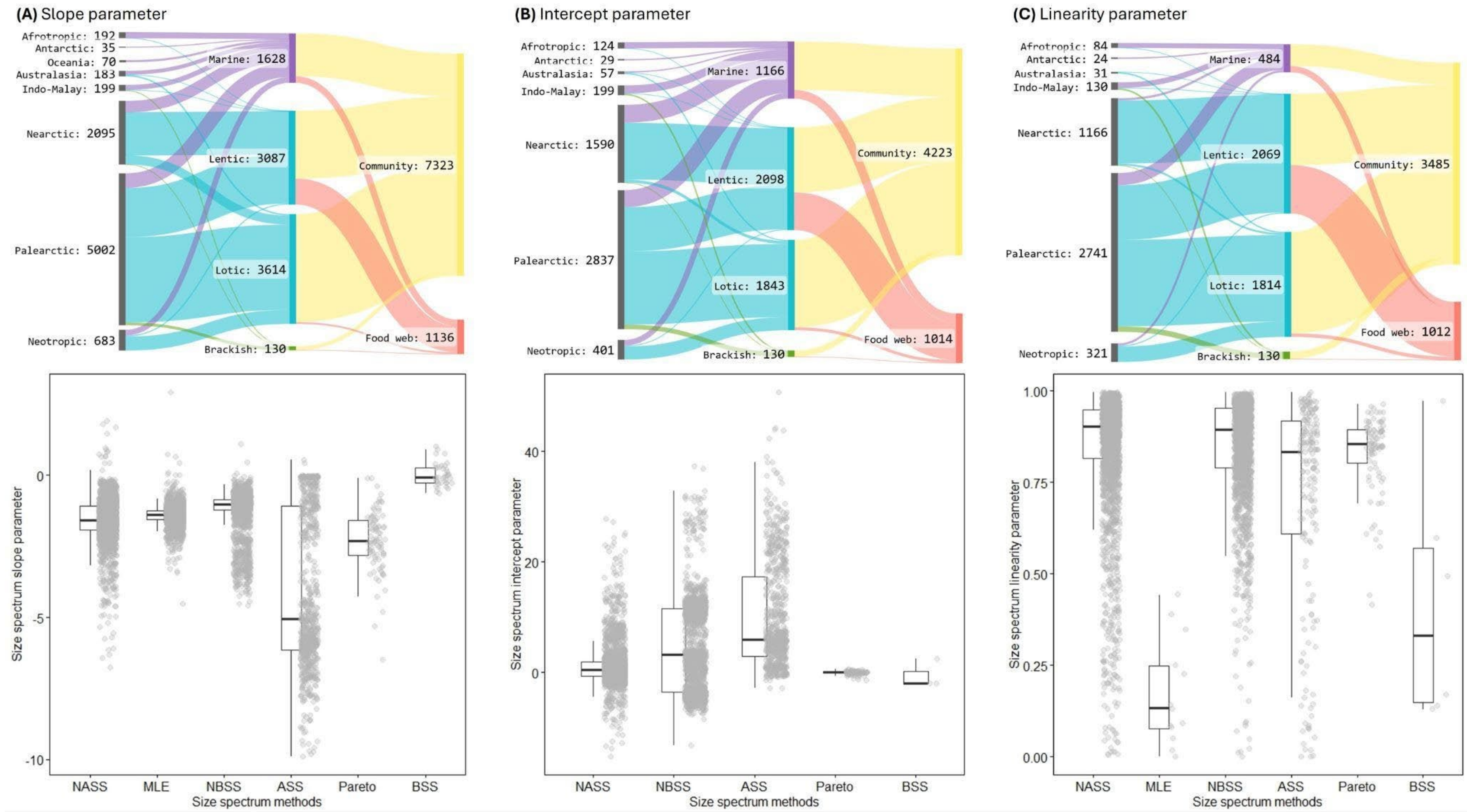
(A) Slope parameter
Afrotropic: 192
Antarctic: 35
Oceania: 70
Australasia: 183
Indo-Malay: 199
Nearctic: 2095
Palearctic: 5002
Neotropic: 683
Marine: 1628
Lentic: 3087
Lotic: 3614
Brackish: 130
Community: 7323
Food web: 1136
(B) Intercept parameter
Afrotropic: 124
Antarctic: 29
Australasia: 57
Indo-Malay: 199
Nearctic: 1590
Palearctic: 2837
Neotropic: 401
Marine: 1166
Lentic: 2098
Lotic: 1843
Brackish: 130
Community: 4223
Food web: 1014
(C) Linearity parameter
Afrotropic: 84
Antarctic: 24
Australasia: 31
Indo-Malay: 130
Nearctic: 1166
Palearctic: 2741
Neotropic: 321
Marine: 484
Lentic: 2069
Lotic: 1814
Brackish: 130
Community: 3485
Food web: 1012
Size spectrum slope parameter
Size spectrum intercept parameter
Size spectrum linearity parameter
NASS
MLE
NBSS
ASS
Pareto
BSS
Size spectrum methods

Figure 4. Alluvial (upper panels) and raincloud (lower panels) plots representing the size spectrum (A) slopes (n = 8,459), (B) intercepts (n = 5,237) and (C) linearity coefficients (n = 4,497) across biogeographic realms, aquatic habitats and levels of ecological complexity (upper panels). For each parameter, lower panels display the raw values according to the different methods used to calculate the size spectrum (method abbreviations are shown in Figure 3). Studies focusing on one broad taxonomic group with similar trophic habits were referred to as "community", while those involving multiple broad taxonomic groups exerting prey-predator dynamics were referred to as "food web". The classification between food webs and community was carried out based on our expert knowledge working in the aquatic taxa. Dots represent individual values for size spectrum parameters, whereas boxplots represent the probability density of the data in aquatic habitats. Alluvial plots were created using open-source, online tool SankeyMATIC (sankeymatic.com).

4. Project personnel:

Zeynep Ersoy. Departamento de Biología y Geología, Física y Química Inorgánica, Universidad Rey Juan Carlos (URJC), Móstoles, Spain.

Charlotte Evangelista. Norwegian Institute for Nature Research (NINA), Trondheim, Norway

Ignasi Arranz. Instituto de Investigación en Cambio Global (IICG-URJC), Universidad Rey Juan Carlos, Tulipán s/n, 28933 Móstoles, España. Departamento de Biología y Geología, Física y Química Inorgánica, Universidad Rey Juan Carlos (URJC), Móstoles, Spain.

Class III. Data set status and accessibility

A. Status

1. Latest update: 13.12.2024.
2. Latest archive date: 13.12.2024.
3. Metadata status: 13.12.2024
4. Data verification: Data were extracted from the literature by the three PIs (Zeynep Ersoy, Charlotte Evangelista and Ignasi Arranz) following the same procedure. We harmonized spatial data by projecting (when necessary) the geographical coordinates using the World Geodetic System (WGS84) as the reference geographic coordinate system. We visually inspected the spatial distribution of the sites to their respective country, region or state boundaries as given in the original data sources. We verified sites with doubtful coordinates using the original data sources. Finally, we checked for typing mistakes and duplicates using preliminary data visualization and analyses in R, and data included in the dataset are, to our best knowledge, free of errors.

B. Accessibility

1. Storage location and medium: The metadata and complete data set are available as Supporting Information at: [Data S1]. Associated data and the R code used for data processing are also available at https://doi.org/10.5281/zenodo.14701391.
2. Contact persons: Charlotte Evangelista (charlotte.evangelista@nina.no); Ignasi Arranz (ignasi.arranz@urjc.es); Zeynep Ersoy (zzeynepersoy@gmail.com). Same address as in Class I.A.
3. Copyright restrictions: The dataset is freely available for non-commercial scientific use (CC BY-NC-SA 4.0 Deed | Attribution-NonCommercial-ShareAlike 4.0 International | Creative Commons).

a. Proprietary restrictions: Please cite this data paper when using its data in publications.

b. Release date: None

c. Citation: Data may be cited following the current Data Paper in Ecology.

d. Disclaimer(s): None

4. Costs: No costs were required to acquiring data.

Class IV. Data structural descriptors

The dataset can be downloaded from the Supporting Information and data repository (Ersoy, Evangelista and Arranz, 2025) and includes four tab-delimited TXT files.

GLOSSAQUA_Dictionary.txt integrates the description of all variables present in the other TXT files. Within GLOSSAQUA_Dictionary.txt, each row corresponds to a variable while the columns contain the name of the variable and its description, an example, and the name of the TXT files where this variable can be found.

GLOSSAQUA_DataSource.txt represents the article and datasets used to collect size spectrum parameters (Table 2). Within GLOSSAQUA_DataSource.txt, each row corresponds to a study while the columns contain the article identification, the reference, the peer-reviewed journal, the year of publication and the Digital Object Identifier (DOI).

GLOSSAQUA_Sample.txt contains information of each sample with its projected geographical coordinates in WGS84 (latitude and longitude; Table 3). Within GLOSSAQUA_Sample.txt, each row corresponds to the sample while columns include the biogeographic realms (using boundaries defined by Olson et al. 2001), habitat and ecosystem types, the context of the study, the level of biological organization, the organismal group and the sampling method.

GLOSAQUA_Size.txt compiles information related to the size spectrum methodology and parameters (Table 4). Within GLOSSAQUA_Size.txt, each row corresponds to the individual estimates of size spectrum parameters, while the columns include the size spectrum parameters (i.e., slope, intercept and linearity), body size units, and body size ranges.

The GLOSSAQUA_DataSource.txt and GLOSSAQUA_Sample.txt files can be merged using the StudyID variable that corresponds to the unique identity of each study (values ranging from 1 to 135). The GLOSSAQUA_Sample.txt and GLOSSAQUA_Size.txt files can be merged using the SampleID variable that indicates the unique identity of each extracted size spectrum (values ranging from 1 to 7943). All data were curated, organized and analyzed using the statistical software R version 4.3.1 (R Core Team 2023). The R code (Figures_GLOSSAQUA_dataset.R) used to import and merge the different datasets, and to display the figures, is available in the data repository (Ersoy, Evangelista and Arranz, 2025).

A. Data set file

1. Identity: GLOSSAQUA_DataDictionary.txt

1a. Size: 5 columns and 46 rows included header row, 10 KB.

1b. Format and storage mode: Plain text (.txt).

1c. Header information: See column descriptions in section B.

1d. Alphanumeric attributes: Mixed.

1e. Special characters/fields: None.

1f. Authentication procedures: None.

2. Identity: GLOSSAQUA_DataSource.txt

2a. Size: 5 columns and 135 rows included header row, 13 KB.

2b. Format and storage mode: Plain text (.txt).

2c. Header information: See column descriptions in section B.

2d. Alphanumeric attributes: Mixed.

2e. Special characters/fields: None.

2f. Authentication procedures: None.

3. Identity: GLOSSAQUA_Sample.txt

3a. Size: 16 columns and 3576 rows included header row, 695 KB.

3b. Format and storage mode: Plain text (.txt).

3c. Header information: See column descriptions in section B.

3d. Alphanumeric attributes: Mixed.

3e. Special characters/fields: None.

3f. Authentication procedures: None.

4. Identity: GLOSSAQUA_Size.txt

4a. Size: 30 columns and 8459 rows included header row, 2 MB.

4b. Format and storage mode: Plain text (.txt).

4c. Header information: See column descriptions in section B.

4d. Alphanumeric attributes: Mixed.

4e. Special characters/fields: None.

4f. Authentication procedures: None.

B. Variable information

Table 2. Source Information in the GLOSSAQUA dataset (GLOSSAQUA_DataSource.txt).

Description of the fields related to the data source.

VariableName	VariableType	Description	Example
StudyID	String	Unique identifier of the data source, either from articles or unpublished individual body size datasets	StudyID_01 (135 unique strings)
Reference	String	Citation of the study	Evans et al. 2022 (135 unique strings)
Journal	String	Names of the journal when data has already been published.	CORAL REEFS (56 unique strings)
PublicationYear	Numeric	Year of publication	2001
DOI	String	Digital Object Identifier (DOI) when data has already been published.	10.1002/lno.11613 (127 unique strings)

Table 3. Site Information in the GLOSSAQUA dataset (GLOSSAQUA_Sample.txt). Description of the fields related to the data site.

VariableName	VariableType	Description	Example
StudyID	String	Unique identifier of the data source, either from articles or unpublished individual body size datasets	StudyID_01 (135 unique strings)
SiteID	String	Unique identifier of the local site	SiteID_01 (3,576 unique strings)
GeographicalLatitude	Numeric	Latitude of the sampling site in decimal degrees	80.76
GeographicalLongitude	Numeric	Longitude of the sampling site in decimal degrees	-179.76
Hemisphere	String	Name of the Hemisphere	"North" and "South"
Realm	String	Classification of the biogeographic realms	"Afrotropic", "Antarctic", "Indo-Malay", "Nearctic", "Neotropic", "Oceania" and "Palearctic"
GeographicalTerritory	String	Name of the main continents, oceans and seas	"Africa", "Antarctic", "Asia", "Europe", "North America", "Oceania", "South America", "Antarctic Ocean", "Arctic Ocean", "Atlantic Ocean", "Indian Ocean", "Mediterranean sea" and "Pacific Ocean"
GeographicalTerritory2	String	Name of the country and sea located within an ocean	Spain
Habitat	String	Type of habitat	"Brackish", "Freshwater" and

			“Marine”
HabitatSpecification	String	Type of habitat	“Brackish”, “Lentic”, “Lotic” and “Marine”
Ecosystem	String	Type of ecosystem	Stream (59 unique strings)
StudyContext	String	Main contexts in which the study was conducted (Table 1)	Natural habitat heterogeneity (19 unique strings)
BiologicalOrganisation	String	Community (e.g., only fish) or Food web (e.g., invertebrate and fish)	“Community” and “Food web”
SpeciesType	String	Name of the organismal group studied (i.e., primary producer, bacteria, protozoan, macroinvertebrate, zooplankton, fish)	Macroinvertebrate+Fish (15 unique strings)
SamplingMethodology	String	General method used to collect data (e.g., electrofishing, acoustic, plankton net, Seine net, Ekman dredge, etc.)	Ekman dredge (44 unique strings)

Table 4. Size spectrum information in the GLOSSAQUA dataset (GLOSSAQUA_Size.txt).

Description of the fields related to the data size spectrum.

VariableName	VariableType	Description	Example
StudyID	String	Unique identifier of the data source, either from articles or unpublished individual body size datasets	StudyID_01 (135 unique strings)
SiteID	String	Unique identifier of the local site	SiteID_01 (3,567 unique strings)
SampleID	String	Unique identifier of the size spectrum (or single entry)	SampleID_1 (8,459 unique strings)
NumSampled	Numeric	The number of times that sites were sampled	377
SamplingYear	Numeric	Year when data was collected	2022
SamplingMonth	String	Month when data was collected. Sometimes multiple months were considered to build the size spectrum	May-Jun-Jul-Aug-Sep Note that "Year" is used when sampling occurred over a year.
NumberSizeSpectrumMethods	Numeric	Number of size spectrum methods used to fit the data	1 or 2
SizeSpectrumMethod	String	Type of size spectrum methodology used to fit the data	"Abundance spectrum (linear)", "Biomass spectrum (linear)", "Maximum Likelihood", "Normalized abundance spectrum (linear)", "Normalized biomass spectrum (linear)" and "Type I Pareto probability density function (power)"
YaxisParameterType	String	Type of variable in the	

		Y-axis	“abundance”, “biomass”, “biomass concentration” and “density”
YaxisParameterTypeSpecification	String	Additional information of the Y-axis regarding if the information explicitly mention that contains carbon or not	“abundance”, “biomass”, “biomass carbon”, “biomass concentration”, “biomass concentration carbon” and “density”
XaxisParameterType	String	Type of body size measured	“body mass” and “physical size”
XaxisParameterTypeSpecification	String	Additional information of the X-axis	“body length”, “body mass”, “body mass carbon” and “body volume”
YaxisParameterUnit	String	Unit used in the Y-axis	g/m2 (49 unique strings)
XaxisParameterUnit	String	Unit used in the X-axis	g (9 unique strings)
XaxisParameterUnitSpecification	String	Additional information of the unit of the X-axis (e.g., dry or wet mass)	g DM and g WM (24 unique strings)
XDimension	String	Dimensionality of the body size	1D or 3D
NumberSizeClasses	Numeric	Number of size classes used to fit the size spectrum model	8
SizeRangeMinimum	Numeric/String	Minimum size used in the first size class	300
SizeRangeMaximum	Numeric/String	Maximum size used in the last size class	2E+27
Slope	Numeric	Value of the size spectrum slope parameter	2.93
SlopeConfIntLow	Numeric	Value of the lower limit of the confident interval of the slope	0.81
SlopeConfIntUp	Numeric	Value of the lower upper of the confident interval of the slope	54.85

SlopeSD	Numeric	Value of the standard deviation of the slope	0.62
SlopeSE	Numeric	Value of the standard error of the slope	4.35
Intercept	Numeric	Value of the size spectrum intercept parameter	50.72
InterceptConfIntLow	Numeric	Value of the lower limit of the confident interval of the intercept	42.66
InterceptConfIntUp	Numeric	Value of the lower upper of the confident interval of the intercept	58.79
InterceptSD	Numeric	Value of the standard deviation of the intercept	5.9
InterceptSE	Numeric	Value of the standard error of the intercept	15.39
Linearity	Numeric	Value of the size spectrum linearity parameter	0.997

1. Variable identity: This information is not available.
2. Variable definition: This information is not available.
3. Units of measurement: This information is not available.
4. Data type
   a. Storage type: Variable structure represents a mix of character and numeric information.
   b. List and definition of variable codes: This information is not available.
   c. Range for numeric values: This information is not available.

d. Missing value codes: Missing values are represented as NAs.

e. Precision: The number of decimals corresponded to a maximum of three.

5. Data format

a. Fixed, variable length: This information is not available.

b. Columns: Start column, end column: This information is not available.

c. Optional number of decimal places: This information is not available.

C. Data anomalies: If no information is available for a given record, this is indicated by NA.

Class V. Supplemental descriptors

A. Data acquisition

1. Data forms or acquisition methods: For studies derived from the literature review, size spectrum parameters were mainly extracted from figures using WebPlotDigitizer software v.3.4 (Rohatgi 2020) but also directly shared by the corresponding authors. Size spectrum parameters for studies derived from the online survey (i.e., individual body size datasets) were calculated using a MLE approach based on a log-likelihood function in the R package sizeSpectra (Edwards 2019; Edwards et al. 2020).
2. Location of completed data forms: This information is not available.
3. Data entry verification procedures: This information is not available.

B. Quality assurance/quality control procedures: The data from individual studies were collected from experts in the fields and raw data (i.e., individual body size) were visually inspected for outliers.

C. Related materials: The references and/or the DOI of each study (i.e., records from the literature review and individual body size datasets from the online survey) are available from the GLOSSAQUA_DataSource.txt file.

D. Computer programs and data-processing algorithms: This information is not available.

E. Archiving

   1. Archival procedures: Data is archived in the folllowing data repository: https://doi.org/10.5281/zenodo.14701391.
   2. Redundant archival sites: We removed duplicated sites from different articles.

F. Publications and results: Results from this dataset are not yet published.

G. History of data set usage

   1. Data request history: None.
   2. Data set update history: None.
   3. Review history: None.
   4. Questions and comments from secondary users: Nearly all biogeographic regions are represented, but it is important to note that our dataset is not free of spatial biases, as many of the samples come from the North Hemisphere, especially the Western Palearctic and Nearctic realms. Thus, researchers should consider geospatial analyses such as spatial autocorrelation in future studies to deal with spatial biases arising from the generation of these data. We also highlight that the size spectrum parameters may not be directly comparable across sites without a full understanding of the sampling methodology, fitting methodology and body size information. This holds in particular for the size spectrum intercept since its values are subjected to the fitting method used (e.g., binned or MLE), the type of variable

used (e.g., biomass or abundance), the sample size and unit of size measurement. To allow comparison among studies, one solution is to standardize (e.g., to have a mean = 0 and a standard deviation = 1) the size spectrum intercepts within studies. We strongly recommend future users to standardize the size spectrum parameters of interest (e.g., see Standardization_GLOSSAQUA_dataset.R code provided in the data repository) before conducting cross-comparison studies. Despite these inherent limitations associated with samples collected for multiple purposes, we are confident that the GLOSSAQUA dataset will stimulate new research in global change ecology and macroecology.

H. Acknowledgments

We thank the authors who kindly shared their data, Jagoba Malumbres-Olarte for the Figure 2 and, two anonymous reviewers for their constructive and useful comments. CE was financially supported by the Government of Catalonia through its Beatriu de Pinós program (E-OR609BP54221). IA is financially supported by the Programa Atracción del Talento (Comunidad de Madrid; project FRESCO: 2022-T1/AMB-24100).

Author contributions

ZE, CE, and IA conceived the idea, acquired funding, built the dataset and wrote the first draft. All authors reviewed, edited and approved the final version of the manuscript.

Literature Citations

Andersen, K. H. 2019. Fish ecology, evolution, and exploitation: a new theoretical synthesis, *Monographs in population biology*. Princeton University Press, Princeton.

Arim, M., Berazategui, M., Barreneche, J. M., Ziegler, L., Zarucki, M., and S. R. Abades. 2011. Determinants of density–body size scaling within food webs and tools for their detection. In *Advances in ecological research* (Vol. 45, pp. 1–39). Academic Press.

Arranz, I., Brucet, S., Bartrons, M., García-Comas, C., and L. Benejam. 2021. Fish size spectra are affected by nutrient concentration and relative abundance of non-native species across streams of the NE Iberian Peninsula. *Science of the Total Environment* **795**: 148792.

Arranz, I., Hsieh, C. H., Mehner, T., and S. Brucet. 2019. Systematic deviations from linear size spectra of lake fish communities are correlated with predator–prey interactions and lake-use intensity. *Oikos* **128**: 33–44.

Arranz, I., Grenouillet, G., and J. Cucherousset. 2023. Human pressures modulate climate-warming-induced changes in size spectra of stream fish communities. *Nature Ecology & Evolution* **7**: 1072–1078.

Atkinson, A., Rossberg, A. G., Gaedke, U., Sprules, G., Heneghan, R. F., Batziakas, S., ... and C. Frangoulis. 2024. Steeper size spectra with decreasing phytoplankton biomass indicate strong trophic amplification and future fish declines. *Nature Communications* **15**: 381.

Basset, A., Barbone, E., Borja, A., Brucet, S., Pinna, M., Quintana, X. D., ... and N. Simboura. 2012. A benthic macroinvertebrate size spectra index for implementing the Water Framework Directive in coastal lagoons in Mediterranean and Black Sea ecoregions. *Ecological Indicators* **12**: 72–83.

Benke, A. C., Huryn, A. D., Smock, L. A., and J. B. Wallace. 1999. Length-mass relationships for freshwater macroinvertebrates in North America with particular reference to the southeastern United States. *Journal of the North American Benthological Society* **18**: 308–343.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and G. B. West. 2004. Toward a metabolic theory of ecology. *Ecology* **85**: 1771–1789.

Chang C-W., Miki T., Shiah F-K., Kao S., Wu J-T., Sastri A. R., and C. Hsieh. 2014. Linking secondary structure of individual size distribution with nonlinear size trophic level relationship in food webs. *Ecology* **95**: 897–909.

Collyer, G., Perkins, D. M., Petsch, D. K., Siqueira, T., and V. Saito. 2023. Land-use intensification systematically alters the size structure of aquatic communities in the Neotropics. *Global Change Biology* **29**: 4094–4106.

Comte, L., Carvajal-Quintero, J., Tedesco, P. A., Giam, X., Brose, U., Erős, T., … and J. D. Olden. 2021. RivFishTIME: A global database of fish time-series to study global change ecology in riverine systems. *Global Ecology and Biogeography* **30**: 38–50.

Deutsch, C., Penn, J. L., Verberk, W. C. E. P., Inomura, K., Endress, M.-G., and J. L. Payne. 2022. Impact of warming on aquatic body sizes explained by metabolic scaling from microbes to macrofauna. *Proceedings of the National Academy of Sciences* **119**: e2201345119.

Dornelas, M., Antão, L. H., Moyes, F., Bates, A. E., … M. L. Zettler. 2018. BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecology and Biogeography* **27**: 760–786.

dos Santos, R. M., Hilbers, J. P., and A. J. Hendriks. 2017. Evaluation of models capacity to predict size spectra parameters in ecosystems under stress. *Ecological Indicators* **79**: 114–121.

Dossena, M., Yvon-Durocher, G., Grey, J., Montoya, J. M., Perkins, D. M., Trimmer, M., and G. Woodward. 2012. Warming alters community size structure and ecosystem functioning. *Proceedings of the Royal Society B* **279**: 3011–3019.

Edwards, A. M., Robinson, J. P., Plank, M. J., Baum, J. K., and J. L. Blanchard. 2017. Testing and recommending methods for fitting size spectra to data. *Methods in Ecology and Evolution* **8**: 57–67.

Edwards, A. M. 2020. sizeSpectra: R package for fitting size spectra to ecological data (including binned data). Available at https://github.com/andrew-edwards/sizeSpectra

Edwards, A. M., Robinson, J. P., Blanchard, J. L., Baum, J., and M. J. Plank. 2020. Accounting for the bin structure of data removes bias when fitting size spectra. *Marine Ecology Progress Series* **636**: 19–33.

Ersoy, Z., Evangelista, C. and Arranz, I. 2025. GLOSSAQUA dataset (v1.0.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14701391

Foo, Y. Z., O'Dea, R. E., Koricheva, J., Nakagawa, S., and M. Lagisz. 2021. A practical guide to question formation, systematic searching and study screening for literature reviews in ecology and evolution. *Methods in Ecology and Evolution* **12**: 1705–1720.

Gaedke, U. 1992. Identifying ecosystem properties: a case study using plankton biomass size distributions. *Ecological Modelling* **63**: 277–298.

Gaedke, U. 1993. Ecosystem analysis based on biomass size distributions: a case study of a plankton community in a large lake. *Limnology and Oceanography* **38**: 112–127.

Gaedke, U., and D. Straile. 1994. Seasonal changes of trophic transfer efficiencies in a plankton food web derived from biomass size distributions and network analysis. *Ecological Modelling* **75**: 435–445.

García-Comas C., Sastri A.R., Ye L., Chang C-Y., Lin F-S., Su M-S., Gong G-C., and C-h. Hsieh. 2016. Prey size diversity hinders biomass trophic transfer and predator size diversity promotes it in planktonic communities. *Proceedings of the Royal Society B* **283**: 20152129.

Ghilarov, M. 1944. Correlation between size and number of soil animals. *Doklady Akademii Nauk* **43**: 283–285.

Gibert J. P., Dell A. I., DeLong J. P., and S. Pawar. 2015. Scaling-up trait variation from individuals to ecosystems. *Advances in Ecological Research* **52**: 1–17.

Hatton, I., Heneghan, R., Bar-On, Y., and E. D. Galbraith. 2021. The global ocean size-spectrum from bacteria to whales. *Science Advances* **7**: eabh3732.

Jennings, S., and J. L. Blanchard. 2004. Fish abundance with no fishing: predictions based on macroecological theory. *Journal of Animal Ecology* **73**: 632–642.

Keddy, P. A., and D. C. Laughlin. 2021. A framework for community ecology: species pools, filters and traits: Cambridge University Press.

Kerr, S. R., and L. M. Dickie. 2001. The biomass spectrum: a predator-prey theory of aquatic production. Columbia University Press, New York, NY.

Kiko, R., Picheral, M., Antoine, D., Babin, M., Berline, L., Biard, T., Boss, E., … and L. Stemmann. 2022. A global marine particle size distribution dataset obtained with the Underwater Vision Profiler 5. *Earth System Science Data* **14**: 4315–4337.

Larrañaga, A., Perkins, D. M., Basaguren, A., Larrañaga, S., Pozo, J., and J. M. Montoya. 2023. Land use drives detritivore size structure and decomposition through shifts in resource quality and quantity. *Science of The Total Environment* 164552.

Liu, K., Jiang, S., Montagnes, D.J.S., Liu, H., Zheng, L., Huang, B., Liu, X., and B. Chen. 2023. Do marine planktonic ciliates follow Bergmann's rule? *Ecography* e06452.

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. McP., Elston, D. A., Scott, E. M., Smith, R. I., Somerfield, P. J., and A. D. Watt. 2010. Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution* **25**: 574–582.

Marin, V., Arranz, I., Grenouillet, G., and J. Cucherousset. 2023. Fish size spectrum as a complementary biomonitoring approach of freshwater ecosystems. *Ecological Indicators* **146**: 109833.

Mehner, T., Attermeyer, K., Brauns, M., Brothers, S., Hilt, S., Scharnweber, K., ... and U. Gaedke. 2022. Trophic transfer efficiency in lakes. *Ecosystems* **25**: 1628–1652.

Méthot, G., C. Hudon, P. Gagnon, B. Pinel-Alloul, A. Armellin and A. M. T. Poirier. 2012. Macroinvertebrate size-mass relationships: how specific should they be? *Freshwater Science* **31**: 750–764.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V., Underwood, E. C., ... and K. R. Kassem. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* **51**: 933–938.

Peralta-Maraver, I., Robertson, A. L., and D. Perkins. 2019. Depth and vertical hydrodynamics constrain the size structure of a lowland streambed community. *Biology Letters* **15**: 20190317.

Perkins, D. M., Durance, I., Jackson, M., Jones, J. I., Lauridsen, R. B., Layer-Dobra, K., Reiss, J., Thompson, M. S. and G. Woodward. 2021. Systematic variation in food web body-size structure linked to external subsidies. *Biology Letters* **17**: 20200798.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., and D. Moher. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* **372**: 71.

Pawar, S., Woodward, G., and A. I.. Dell. 2015. Trait-Based Ecology - from structure to function. Academic Press.

Petchey, O. L., and A. Belgrano. 2010. Body-size distributions and size-spectra: universal indicators of ecological status? *Biology Letters* **6**: 434–437.

Peters, R. H. 1983. The ecological implications of body size. Cambridge University Press, Cambridge.

Platt, T., and K. Denman. 1977. Organisation in the pelagic ecosystem. *Helgolander Wiss. Meeresunters* **30**: 575–581.

Pomeranz, J. P. F., Junker, J. R., and J. S. Wesner. 2022. Individual size distributions across North American streams vary with local temperature. *Global Change Biology* **28**: 848–858.

Potapov, A. M., Rozanova, O. L., Semenina, E. E., Leonov, V. D., Belyakova, O. I., Bogatyreva, V. Y., Degtyarev, M. I., Esaulov, A. S., Korotkevich, A. Y., Kudrin, A. A., Malysheva, E. A., Mazei, Y. A., Tsurikov, S. M., Zuev, A. G., and A. V. Tiunov. 2021. Size compartmentalization of energy channeling in terrestrial belowground food webs. *Ecology* **102**: e03421.

Rohatgi, A. 2020. Webplotdigitizer: Version 4.3. 2020. https://automeris.io/WebPlotDigitizer.

R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sheldon, R. W., Prakash, A., and W. H. Sutcliffe. 1972. The size distribution of particles in the ocean. *Limnology and Oceanography* **17**: 327–340.

Silvert, W., and T. Platt. 1978. Energy flux in the pelagic ecosystem: A time-dependent equation. *Limnology and Oceanography* **23**: 813–816.

Sprules, W. G. and M. Munawar. 1986. Plankton size in relation to ecosystem productivity, size, and perturbation. *Canadian Journal of Fisheries and Aquatic Sciences* **43**: 1789–1794.

Sprules, W. G., and L. E. Barth. 2016. Surfing the biomass size spectrum: some remarks on history, theory, and application. *Canadian Journal of Fisheries and Aquatic Sciences* **73**: 477–495.

Sprules, W. G. 2022. The Biomass Size Spectrum, in: Encyclopedia of Inland Waters. Elsevier, pp. 390–400.

Tomanova, S., Poulet, N., Coullier, D., and M. Thiret. 2010. Relations longueur/poids pour les poissons d'eau douce en France Rédaction S. Office National de l´Eau et des Milleux Aquatiques.

U.S. Environmental Protection Agency. 2010. Standard Operating Procedure for Phytoplankton Analysis. LG401, Revision 05, February 2010.

Vidondo, B., Prairie, Y. T., Blanco, J. M., and C. M. Duarte. 1997. Some aspects of the analysis of size spectra in aquatic ecology. *Limnology and oceanography* **42**: 184–192.

White, E. P., Ernest, S. M., Kerkhoff, A. J., and B. J. Enquist. 2007. Relationships between body size and abundance in ecology. *Trends in Ecology & Evolution* **22**: 323–330.

Woodward, G., Ebenman, B., Emmerson, M., Montoya, J., Olesen, J., Valido, A., and P. Warren. 2005. Body size in ecological networks. *Trends in Ecology & Evolution* **20**: 402–409.