

scientific data



OPEN
ARTICLE

CRITTERBASE, a science-driven data warehouse for marine biota

Katharina Teschke^{1,2,8}✉, Casper Kraan^{1,2,7,8}, Paul Kloss^{1,2}, Henrike Andresen^{1,2,7}, Jan Beermann^{1,2}, Dario Fiorentino^{1,2,7}, Manuela Gusky^{1,2}, Miriam L. S. Hansen³, Rebecca Konijnenberg¹, Roland Koppe¹, Hendrik Pehlke^{1,2}, Dieter Piepenburg^{1,2,3}, Tawfik Sabbagh^{1,4}, Alexa Wrede^{1,2,5}, Thomas Brey^{1,2,6} & Jennifer Dannheim^{1,2}

Data on marine biota exist in many formats and sources, such as published literature, data repositories, and unpublished material. Due to this heterogeneity, information is difficult to find, access and combine, severely impeding its reuse for further scientific analysis and its long-term availability for future generations. To address this challenge, we present CRITTERBASE, a publicly accessible data warehouse and interactive portal that currently hosts quality-controlled and taxonomically standardized presence/absence, abundance, and biomass data for 18,644 samples and 3,664 benthic taxa (2,824 of which at species level). These samples were collected by grabs, underwater imaging or trawls in Arctic, North Sea and Antarctic regions between the years 1800 and 2014. Data were collated from literature, unpublished data, own research and online repositories. All metadata and links to primary sources are included. We envision CRITTERBASE becoming a valuable and continuously expanding tool for a wide range of usages, such as studies of spatio-temporal biodiversity patterns, impacts and risks of climate change or the evidence-based design of marine protection policies.

Introduction

Marine ecosystems provide major functions and services, which originate from the interactions of many biota a particular community consists of. However, more often than not marine biogeographical research^{1,2} and large-scale ecosystem management plans have no access to suitable community data sets for a number of reasons: (1) with increasing water depth, samples are increasingly more difficult and expensive to collect. (2) The deeper the samples are taken, the less is known about the taxonomic provenance. (3) Sorting of samples, especially of invertebrates, is time-consuming and expensive. (4) Identifying specimens requires profound taxonomic expertise, which is generally becoming rarer in the scientific community. As a result, (5) spatial resolution and coverage of samples are often limited, prohibiting large-scale analyses³. Moreover, (6) merging of data sets from different sources involves time-consuming synchronization of taxonomic information. Finally, (7) many data sets are not open-access and exist only in spreadsheets and local databases. These factors severely impede the scientific analysis and reuse of such data. Biogeographical research generally relies on the use of publicly available data to study large-scale biodiversity patterns in relation to environmental change, human impacts, or to be able to take protective measures⁴. Consequently, appropriate data management tools that comply with the FAIR principles (Findable, Accessible, Interoperable, Reusable) are needed to achieve an effective data and knowledge transfer to support scientific advice to decision-makers and stakeholders.

We introduce the open-access data-warehouse CRITTERBASE (<https://critterbase.awi.de>) for marine biota that intends to remedy these issues and facilitate functional biogeographic studies and ecosystem management approaches at multiple spatial scales. The development and implementation of CRITTERBASE focused on benthic data so far, because this is our main field of expertise and here we see the most urgent need to act.

¹Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12, 27570, Bremerhaven, Germany. ²Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg, Ammerländer Heerstraße 231, 23129, Oldenburg, Germany. ³University of Kiel, Institute for Ecosystem Research, Olshausenstraße 75, 24118, Kiel, Germany. ⁴Hochschule Bremerhaven, An der Karlstadt 8, 27568, Bremerhaven, Germany. ⁵Helmholtz-Zentrum Hereon, Institute of Carbon Cycles, Max-Planck-Straße 1, 21502, Geesthacht, Germany. ⁶University of Bremen, Bibliothekstraße 1, 28359, Bremen, Germany. ⁷Present address: Thünen Institute for Sea Fisheries, Herwigstraße 31, 27572, Bremerhaven, Germany. ⁸These authors contributed equally: Katharina Teschke, Casper Kraan. ✉e-mail: Katharina.Teschke@awi.de

Source	Area	Period	<i>n</i> -samples	Sampling method	Taxonomic resolution
PANABIO	Arctic seas	1800–2014	11,575	Photographs, various trawling and grab methods	Species, genus
BENOSIS	North Sea	1969–2007	6,435	Various grabs, box cores, otter trawls and beam trawls	Species
WEECOS	Weddell Sea	1983–2004	634	Various grabs, trawls, traps and photographs	Species to phylum

Table 1. Summary of CRITTERBASE contents.

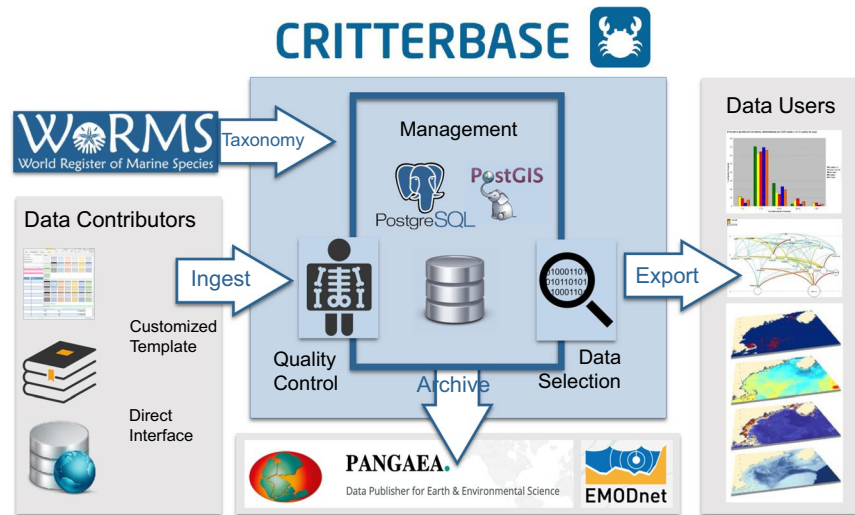


Fig. 1 Workflow of CRITTERBASE. Ingesting data into the data-warehouse and generating output for data visualization and extraction.

Benthic communities play a key role in processes of important functions and services of marine ecosystems^{5,6}. They contribute significantly to overall marine biodiversity, constitute important parts of marine food webs, and facilitate nutrient recycling at the sediment-water interface⁷. Moreover, benthic organisms are excellent sentinels of changing environmental conditions⁸, such as eutrophication⁹, owing to the close association between environmental drivers and benthic distribution patterns^{10,11} and to the comparatively long lifespans of benthic organisms. Benthic biodiversity data that cover large spatial scales and long temporal scales at high taxonomic resolution are therefore of pivotal importance for marine ecosystem management and environmental protection to ensure a sustainable use of coastal and offshore systems. Key examples of this are management approaches set out in national and international guidelines, such as the Marine Strategy Framework Directive (MSFD) of the European Union, or marine spatial planning focusing on the key functions of ecosystems¹².

Results

Application state. Currently, CRITTERBASE hosts data on benthos from Arctic, North Sea, and Antarctic regions (see Table 1), as these are the geographic foci of our research at the Alfred Wegener Institute (AWI). CRITTERBASE uses a shared data model for data from all geographic regions that safeguards the integrity of data regardless of whether they are collated from literature¹³, cooperating researchers, own research, archives or repositories¹⁴. To this end, it utilizes a single standardized workflow (Fig. 1). Only data with metadata on sampling location and date, taxonomic resolution, and sampling method were included, leading to the spatial extent of current records shown in Fig. 2. Data quality controls are major components within CRITTERBASE and ensure that the imported data meet a high quality standard. There are basic quality components, such as its data model itself, and several other routines that flag mistakes through a number of logical checks before, during and after data import to prevent data errors that may corrupt subsequent analyses. Further details on the quality control components are provided at <https://critterbase.awi.de/#qc>.

Data provenance. The marine benthic data collated originated from three regional projects: PANABIO - PAN-Arctic Information System of Benthic BIOta (Arctic), BENOSIS - Benthic North Sea Information System (North Sea), and WEECOS - WEddell Sea Integrated ECOSystem (Antarctic).

The majority of the Arctic data originated from a pan-Arctic inventory of Arctic benthic fauna¹³ as part of the Census of Marine Life's Arctic Ocean Diversity project, or had been published by the data publisher PANGAEA (www.pangaea.de), e.g. Andrade *et al.*¹⁴. Most records were identified to species level, but some to genus level only. Users can find trait information of Arctic benthos in the open-access Arctic Traits Database¹⁵, which, like CRITTERBASE, also uses the World Register of Marine Species (WoRMS; www.marinespecies.org) to assign taxonomic identities.

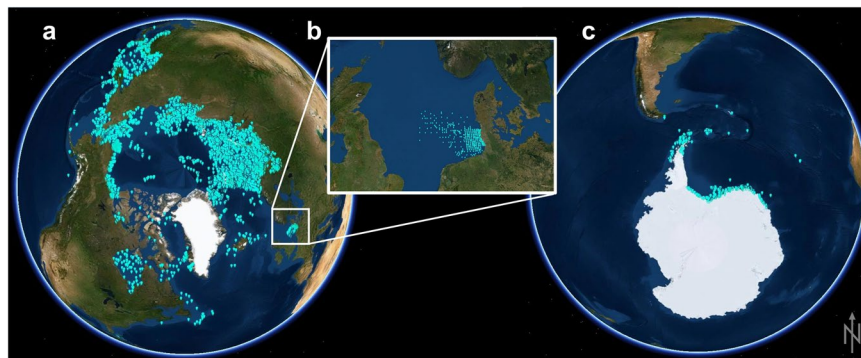


Fig. 2 Map of 14,212 sampling stations in (A) Arctic seas (11,561 stations), (B) the North Sea (2,119 stations), and (C) the Weddell Sea and adjacent seas (Antarctica) (532 stations).

The North Sea benthic data are a synthesis of 13 projects. Data include grab and trawl samples taken in the southeastern North Sea between 1969 and 2007. The organisms have been identified to the lowest taxonomic level possible, i.e. the majority of data are available at species level. Data from four projects have been published^{16–19}, while the remaining nine data sets will be made available in CRITTERBASE in due course.

Antarctic benthic data^{20–24} were compiled, processed and uploaded to PANGAEA as part of the Weddell Sea marine protected area planning project^{25,26}, with the exception of data already published by Gerdes *et al.*^{27–41}. The data mainly include trawls and grab samples taken in the Weddell Sea and adjacent seas between the early 1980s and mid-2000s (Table 1). Most data are at species level and have a particular focus on echinoderm taxa. Two data collections encompass only higher taxonomic groups (mostly at the level of order or higher).

Data query. CRITTERBASE is an open-access information system, allowing users to query and download data as Excel files for further processing. Users are kindly asked to cite this paper in any publication or report and provide attribution to the original data sources that form the basis of CRITTERBASE. CRITTERBASE's query functions allow for the selection of data records based on (1) location (structured according to the regions used by the Food and Agriculture Organization of the United Nations, FAO), (2) sampling gear (e.g., trawl, grab, video) and (3) data set. Each data record is stored as part of a unique data set, which is accessible via a unique identifier (DOI or reference to the original publication containing the data) and generally represents a cruise or a research project. Data records represent the occurrence of a specific taxon at a given geographic position at a given time. Additional information on the abundance, biomass and life stage of the taxa are provided for 25%, 20% and 0.6% of all data sets respectively. Each data record has a unique identification number within CRITTERBASE. Taxonomic information is recorded using AphiaIDs of WoRMS that allow for easy scaling from species to higher taxonomic level.

Discussion

The main objective of CRITTERBASE is to promote the sharing of and easy access to data for the marine biology community, serving as a tool for collaboration while also safeguarding valuable data for future use. The data can be used, for example: (i) to support impact studies (e.g., effects of trawling from commercial fishing or the influence of offshore wind farms), (ii) as a knowledge base to be used for marine spatial planning and associated management and monitoring activities (e.g., within the framework of the MSFD, the EU Habitats Directive, or the International Council for the Exploration of the Sea [ICES]), (iii) to assess trends of benthic biodiversity over space and time⁴², such as species' range shifts or the introduction of invasive species due to climate change^{43–45}.

CRITTERBASE provides a valuable resource for research in polar and temperate regions, as well as for scientists aiming to address large-scale patterns of biodiversity. However, there are a few caveats that need to be considered due to the large variety of sampling methods and taxonomic resolution. Benthos samples were collected on board of research vessels with commonly used methods, such as trawling, grab sampling, or recorded via seabed imaging by means of ocean-floor observation systems and variations thereof¹³. The sampling methods cover different, yet overlapping, parts of the benthic communities. Grab samples primarily contain macrobenthic fauna, trawl catches encompass megabenthos, and seabed images give information on epibenthos. This difference in species-scope is noted in the associated metadata and should be considered when querying or analyzing data. Sampling depths range from coastal waters to areas on the continental shelf and the deep sea (specifically, from a few meters to more than 4000 m). Samples were processed using standard operating procedures, such as the ICES 1999 ISO standard for grab sampling. However, it should be noted that this was not done homogeneously across areas and is not standardized across data sets. Metadata detail each entry's methodological variations, and, where available, a reference is provided for each data set to provide more detail on the applied methodologies and their limitations. For example, some research cruises were purely focused on certain taxonomic groups⁴⁶, thus restricting the taxonomic scope of the data. This is also noted in the metadata and should be carefully reviewed while querying or processing the data. Taxonomic adjustments of the data might be necessary depending on the user's needs. For instance, it may be necessary to aggregate species-level taxa to higher taxonomic groups when merging data sets with different taxonomic resolution. It should also be noted that the records currently within CRITTERBASE indicate only the presence of a recorded taxon. While

absences can be inferred from the taxonomic scope of each data set, this is a decision that has to be made by the data analyst and is case-dependent.

In summary, CRITTERBASE's extensive metadata (from information on catchability, such as mesh size or trawling speed, to details on the taxonomic resolution and coverage) provides the best possible transparency of the data. This, in turn, allows users to perform analyses on joint data sets and thereby contributes to a better understanding across large temporal and spatial scales. It is important to note, however, that case-by-case decisions are necessary to appropriately pool or aggregate the partly inconsistent data sets (i.e., data sets differing in sampling approaches) provided by CRITTERBASE so that the data can eventually be used for joint analyses⁴⁷. Data analysts should pay adequate attention and be vigilant in order to correctly merge the data for their purpose before starting any data analysis. Negligent use of the data provided by CRITTERBASE could lead to wrong ecological conclusions regarding, for instance, temporal and spatial trends of biodiversity.

Future releases of CRITTERBASE will contain more software features and will include data on more diverse biota. We aim to maintain CRITTERBASE as a common open access platform. We plan to grow the platform by including data from other researchers or groups (e.g. ICES WGM BRED data initiative). While the current version of CRITTERBASE includes benthic invertebrates only, future developments of the data model will allow for the ingestion of data from other marine realms, such as pelagic and under-ice invertebrates and fish. CRITTERBASE will also be upgraded to handle information on functional traits and further organism-specific properties⁴⁸. This will then encourage and facilitate trait-based approaches (alongside species-based methods) to, for example, improve our understanding of the processes that influence biodiversity-ecosystem functioning⁹ and to assess trends in ecosystem functioning over space and time. In addition, we will extend the CRITTERBASE platform with routines (implemented in R or Python) for the calculation of metrics useful for further applications, such as estimating secondary production or carrying out effective survey planning.

Methods

Data compilation. The compilation of data included in this release of CRITTERBASE consisted of two initial steps. First, we defined a data model for CRITTERBASE that would serve Arctic, North Sea, and Antarctic data demands for current and future use (for details on the data model, see <https://critterbase.awi.de/#qc#critterbase-data-model>).

For each geographic area, efforts had already been made to build a data-warehouse that would support species distribution modeling (e.g., PANABIO¹³ in the Arctic seas), human-impact studies (e.g., benthos database for ecological research in the North Sea: BENOSIS), or marine conservation planning in the Weddell Sea^{25,26} (WEECOS). These previous efforts were merged into a general data model through iterative discussions with the data warehouse developers. We then ingested already available data into CRITTERBASE (Table 1), including those available from PANGAEA¹⁴ and our own research. The data were then quality checked by AWI experts in data collection and taxonomic identification of benthic communities.

As a result of the data compilation process, CRITTERBASE's initial data sets reflect the research needs of the department of Functional Ecology within AWI. They represent a first stock of benthic data and serve to demonstrate and develop the benefits of using CRITTERBASE when compared to traditional spreadsheet-based archives. The open-access CRITTERBASE web interface allows any user to query and download data for further processing for any purpose.

Data quality control. The automated quality management procedures built into the CRITTERBASE Collector App - the backend of the CRITTERBASE web interface - prevent the import of incorrect or incomplete data sets. Common mistakes are identified through a number of logical controls before, during and after data import. These include, for instance, the detection of differing sampling dates, coordinates or water depths within a single sample. In addition, any new taxonomic name not previously imported via a data set is validated against the current taxonomic classification provided by WoRMS. This prevents the use of synonyms, any incorrect spelling of scientific names, or the use of outdated names, all of which potentially inflate biodiversity estimates and skews species distributions. We also implemented checkpoints in the CRITTERBASE Collector App to verify the type of data we are dealing with (e.g., presence/absence versus abundance data), which ultimately decides on the kinds of biodiversity analysis possible (Fig. 3). Two types of error messages are possible during data ingestion: a complete rejection of the data due to critical errors that comprise the integrity of the database and need to be solved immediately, or minor warnings indicating mistakes that could be improved to increase the data quality, such as small differences in the spelling of scientific names (see details on CRITTERBASE quality management and its control components at <https://critterbase.awi.de/#qc>).

In addition, the CRITTERBASE Collector App, available at Zenodo (<https://doi.org/10.5281/zenodo.5724021>⁴⁹), enables users to create their own CRITTERBASE to manage ecological data projects on their own devices. It sets up a clean PostgreSQL object-relational database with the spatial database extender PostGIS and allows users to quality-check and store biological data (using the CRITTERBASE data model) without having to share their data through the open-access CRITTERBASE web interface (<https://critterbase.awi.de>). Users working locally can make direct queries to their CRITTERBASE via SQL, R and Python, allowing users to keep data queries and code neatly in one place for subsequent analyses. We hope that the option for decentralized working with the CRITTERBASE Collector App - as an open-source tool - will make this an appealing data management option for other researchers, resulting in more processed and quality-checked data sets, which in turn could be made available for publication via the CRITTERBASE web interface.

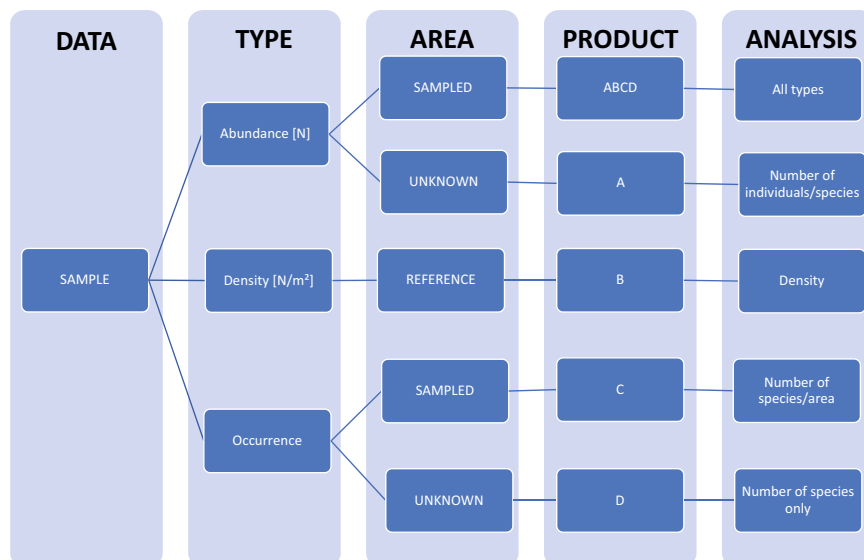


Fig. 3 Flow chart indicating the various data scenarios for biota and sampling area input in CRITTERBASE. Starting from taking a sample, biotic data can be abundances, densities (i.e. in our case individuals/m²) or occurrences. Depending on the availability of sampling area, i.e. actually sampled area, reference area (fictional/defined area; here: 1 m²) or unknown area, possible data product types are ABCD, A, B, C, or D. A allows to assess the relationship between number of individuals and number of species per sample, B allows for the assessment of density, C only allows for the determination of the relationship between number of species and area, whereas D only allows for the determination of the number of species per sample. Type ABCD supports all types of biodiversity analysis. An incorrect pairing of Type and Area (e.g., abundance with reference area) produces an error message upon attempted import into CRITTERBASE. If there is actually no information on the sampled area, unknown area must be selected (which is accepted and leads to Product A). If information on the reference area (instead of sampled area) is given, abundance values would have to be entered as densities. This quality check ensures that the data type and area are consistent.

Data availability

All data from Arctic, North Sea, and Antarctic regions, discussed in this paper, are publicly available at <https://critterbase.awi.de>. Continuous online access to CRITTERBASE will be ensured through a permanent data portal hosted by the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (<https://marine-data.org>). New data will be published as they are submitted and meet the quality standards.

Code availability

The source code of the web service and portal (CRITTERBASE Version 1.0.0) - for accessing quality-checked data sets - is released under the BSD 3 license and is publicly available via the following stable link: <https://critterbase.awi.de/code>. The software libraries and versions used are referenced in the source code.

The Collector App for CRITTERBASE - the software for data quality checking and storage - is available as a free download under the open-source BSD 3 license (<https://doi.org/10.5281/zenodo.5724021>⁴⁹). The software libraries and versions used are referenced in the README.MD file.

Received: 8 December 2021; Accepted: 25 July 2022;

Published online: 06 August 2022

References

- Violle, C., Reich, P. B., Pacala, S. W., Enquist, B. J. & Kattge, J. The emergence and promise of functional biogeography. *Proc. Nat. Acad. Sci.* **111**, 13690–13696 (2014).
- Robinson, L. M. *et al.* Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. *Global Ecol. Biogeogr.* **20**, 789–802 (2011).
- Bijleveld, A. I. *et al.* Designing a benthic monitoring programme with multiple conflicting objectives. *Methods Ecol. Evol.* **3**, 526–536 (2012).
- Costello, M. J., Horton, T. & Kroh, A. Sustainable biodiversity databasing: International, collaborative, dynamic, centralised. *Trends Ecol. Evol.* **33**, 803–805 (2018).
- Pecl, G. T. *et al.* Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* **355**, eaai9214 (2017).
- Jørgensen, L. L. *et al.* Impact of multiple stressors on sea bed fauna in a warming Arctic. *Mar. Ecol. Prog. Ser.* **608**, 1–12 (2019).
- Lohrer, A. M., Thrush, S. F., Hewitt, J. E. & Kraan, C. The up-scaling of ecosystem functions in a heterogeneous world. *Sci. Rep.* **5**, 10349 (2015).
- Laggen, C. *et al.* Climate change, glacier retreat and a new ice-free island offer new insights on Antarctic benthic responses. *Ecography* **41**, 579–591 (2018).
- Thrush, S. F. *et al.* Changes in the location of biodiversity-ecosystem function hot spots across the seafloor landscape with increasing sediment nutrient loading. *Proc. Royal Soc. London B: Biol. Sci.* **284**, 20162861 (2017).

10. Kraan, C., Aarts, G., van der Meer, J. & Piersma, T. The role of environmental variables in structuring landscape-scale species distributions in seafloor habitats. *Ecology* **91**, 1583–1590 (2010).
11. Kraan, C., Dormann, C. F., Greenfield, B. L. & Thrush, S. F. Cross-scale variation in biodiversity-environment links illustrated by coastal sandflat communities. *PLoS ONE* **10**, e0142411 (2015).
12. Ingeman, K.E., Samhuri, J.F., Stier, A.C. Ocean recoveries for tomorrow's Earth: Hitting a moving target. *Science* **363** (2019).
13. Piepenburg, D. *et al.* Towards a pan-Arctic inventory of the species diversity of the macro- and megabenthic fauna of the Arctic shelf seas. *Mar. Biodiv.* **41**, 51–70 (2011).
14. Andrade, H. *et al.* Benthic fauna in soft sediments from the Barents and Pechora Seas. *PANGAEA* <https://doi.org/10.1594/PANGAEA.877932> (2017).
15. Degen, R. & Faulwetter, S. The Arctic Traits Database – A repository of arctic benthic invertebrate traits. *Earth Syst. Sci. Data* **11**, 301–322 (2019).
16. Armonies, W. Macrozoobenthos and sediment composition in the North Sea and the Wadden Sea near the Island of Sylt between 1992 and 2002. *PANGAEA* <https://doi.org/10.1594/PANGAEA.875690> (2017).
17. Rachor, E. & Nehmer, P. Abundance of benthic infauna in surface sediments from the North Sea sampled during HEINCKE cruise HE133. *PANGAEA* <https://doi.org/10.1594/PANGAEA.756768> (2003).
18. Schröder, A. *et al.* Benthic fauna at station FINO 1, 2005–2007. *PANGAEA* <https://doi.org/10.1594/PANGAEA.805200> (2008).
19. Schröder, A. & Rachor, E. Abundance and biomass of in- and epibenthic invertebrate macrofauna in the German Bight from 1969 to 2000. *PANGAEA* <https://doi.org/10.1594/PANGAEA.667646> (2007).
20. Teschke, K. & Brey, T. Presence and absence records of benthic taxa from trawl samples taken in the Weddell Sea (Antarctica) and neighbouring seas during “Polarstern” cruises ANT VII/4, ANT IX/3, ANT XIII/3, ANT XV/3, ANT XVII/3 and ANT XXI/2 between 1989 and 2004. *PANGAEA* <https://doi.org/10.1594/PANGAEA.899085> (2019).
21. Teschke, K. & Brey, T. Presence and absence records of sea star species (class: Asteroidea) from trawl, grab and trap samples in the Weddell Sea and western Antarctic Peninsula region during “Polarstern” cruises ANT I/2, ANT II/4, ANT V/3, ANT VI/3, ANT XV/3 and ANT XVII/3. *PANGAEA* <https://doi.org/10.1594/PANGAEA.898629> (2019).
22. Teschke, K. & Brey, T. Abundance records of five most abundant brittle star species (Ophiuroidea) from trawl, grab and trap samples in the Weddell Sea and neighbouring seas during “Polarstern” cruises ANT I/2, ANT II/4, ANT V/3 and 4, ANT VI/3, ANT IX/3 and ANT X/3 between 1983 and 1992. *PANGAEA* <https://doi.org/10.1594/PANGAEA.898773> (2019).
23. Teschke, K. & Brey, T. Abundance records of polychaete taxa (family: Aphroditidae, Polynoids) from trawl samples in the Weddell Sea and neighbouring seas during “Polarstern” cruises ANT I/2, ANT II/4, ANT V/3 and 4, ANT VII/4, ANT IX/3 and ANT X/3 between 1983 and 1992. *PANGAEA* <https://doi.org/10.1594/PANGAEA.902858> (2019).
24. Teschke, K., Gutt, J. & Brey, T. Abundance and biomass records of two sea urchin species (*Sterechinus antarcticus*, *S. neumayeri*) from trawls, grabs, traps and photographs in the Weddell Sea during POLARSTERN cruises ANT-I/2, ANT-II/4, ANT-III/3, ANT-V/3-4, ANT-VI/3 and ANT-VII/4. *PANGAEA* <https://doi.org/10.1594/PANGAEA.898706> (2019).
25. Teschke, K. *et al.* An integrated compilation of data sources for the development of a marine protected area in the Weddell Sea. *Earth Syst. Sci. Data* **12**, 1003–1023 (2020).
26. Teschke, K. *et al.* Planning marine protected areas under the CCAMLR regime – The case of the Weddell Sea (Antarctica). *Mar. Policy* **124** (2021).
27. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-XXVII/3. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834057> (2014).
28. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-XXI/2. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834049> (2014).
29. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-XIX/5. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834069> (2014).
30. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-XVII/3 (EASIZ III). *PANGAEA* <https://doi.org/10.1594/PANGAEA.834074> (2014).
31. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-XV/3. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834041> (2014).
32. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-X/3. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834025> (2014).
33. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-IX/3. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834013> (2014).
34. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-VII/4. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834021> (2014).
35. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-V/1. *PANGAEA* <https://doi.org/10.1594/PANGAEA.717708> (2014).
36. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-XIII/4. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834033> (2014).
37. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-XXIII/8. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834053> (2014).
38. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-VI/3. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834017> (2014).
39. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-III/2. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834009> (2014).
40. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during Walther Herwig II cruise WH068/1. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834061> (2014).
41. Gerdes, D. Abundance of macrozoobenthos in surface sediments sampled during Walther Herwig II cruise WH068/2. *PANGAEA* <https://doi.org/10.1594/PANGAEA.834065> (2014).
42. McGill, B. J., Dornelas, M., Gotelli, N. J. & Magurran, A. E. Fifteen forms of biodiversity trend in the Anthropocene. *Trends Ecol. Evol.* **30**, 104–113 (2015).
43. Goldsmit, J. *et al.* Projecting present and future habitat suitability of ship-mediated aquatic invasive species in the Canadian Arctic. *Biol. Inv.* **20**, 501–517 (2018).
44. Sorte, C. J. B., Williams, S. L. & Carlton, J. T. Marine range shifts and species introductions: comparative spread rates and community impacts. *Global Ecol. Biogeogr.* **19**, 303–316 (2010).
45. Walther, G.-R. *et al.* Alien species in a warmer world: risks and opportunities. *Trends Ecol. Evol.* **24**, 686–693 (2009).
46. Bilyard, G. R. & Carey, A. G. Distribution of western Beaufort Sea polychaetous annelids. *Mar. Biol.* **54**, 329–339 (1979).
47. Jørgensen, L. L. *et al.* International megabenthic long-term monitoring of a changing arctic ecosystem: Baseline results. *Prog. Oceanogr.* **200**, 102712 (2022).
48. Degen, R. *et al.* Trait-based approaches in rapidly changing ecosystems: A roadmap to the future polar oceans. *Ecol. Ind.* **91**, 722–736 (2018).
49. Kloss, P. *et al.* The Collector's App for CRITTERBASE, a science-driven data warehouse for marine biota, *Zenodo* <https://doi.org/10.5281/zenodo.5724021> (2021).

Acknowledgements

We acknowledge the efforts of all collaborators to make their data available in CRITTERBASE. HIFMB is a collaboration between the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, and the Carl-von-Ossietzky University Oldenburg, initially funded by the Ministry for Science and Culture of Lower Saxony and the Volkswagen Foundation through the “Niedersächsisches Vorab” grant program (grant number ZN3285). K. Teschke, R. Konijnenberg, H. Pehlke and P. Kloss were financially supported by the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office for Agriculture and Food (BLE) (grant number 2819HS015). J. Dannheim, H. Pehlke and P. Kloss were funded by the German Federal Ministry for Economic Affairs and Energy (grant number 0325921). M. Hansen was financially supported by the German Federal Ministry of Education and Research (grant number 03F0776). J. Beermann was funded by the German Federal Agency for Nature Conservation (grant number 3519532201). CRITTERBASE is a use case in the DFG project “National Research Data Initiative for Biodiversity NFDI4Biodiversity”. We thank two anonymous reviewers for constructive comments on the manuscript.

Author contributions

K.T., C.K., M.L.S.H., J.D., H.P., J.B., A.W., D.P. collated the data. P.K. developed the software with input from K.T., C.K., T.B., D.P., J.D., H.P., J.B., H.A., D.F., M.G., C.K. and K.T. wrote the manuscript with contributions from all co-authors.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.T.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022