Title

The Antarctic Seafloor Annotated Imagery Database

Authors

Jan Jansen[1], Victor Shelamoff[1], Charley Gros[1], Thomas Windsor[1], Nicole A. Hill[1], David K. Barnes[2], David A. Bowden[3], Julian Gutt[4], Narissa Bax[5,6], Rachel Downey[7], Marc Eléaume[8], Alexandra L. Post[9], Huw Griffiths[2], Katrin Linse[2], Dieter Piepenburg[4], Autun Purser[4], Craig R. Smith[10], Amanda F. Ziegler[11] & Craig R. Johnson[1]

Affiliations

1. Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, AUS
2. British Antarctic Survey, UKRI, Cambridge, UK
3. National Institute of Water and Atmospheric Research, Wellington, NZ
4. Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, GER
5. South Atlantic Environmental Research Institute, Stanley, FLK
6. Centre for Marine Socioecology, Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, AUS
7. Fenner School of Environment and Society, Australia National University, Canberra, AUS
8. Muséum National d'Histoire Naturelle, UMR 7205-ISYEB, Centre National de la Recherche Scientifique-UPMC-EPHE, Paris, FRA
9. Geoscience Australia, Canberra, AUS
10. Department of Oceanography, University of Hawai'i at Mānoa, Honolulu, USA
11. Department of Arctic and Marine Biology, The Arctic University of Norway, Tromsø, NOR

corresponding author: Jan Jansen (jan.jansen@utas.edu.au)

Abstract

Marine imagery is a comparatively cost-effective way to collect data on seafloor organisms, biodiversity and habitat morphology. However, annotating these images to extract detailed biological information is time-consuming and expensive, and reference libraries of consistently annotated seafloor images are rarely publicly available. Here, we present the Antarctic Seafloor Annotated Imagery Database (AS-AID), a result of a multinational collaboration to collate and annotate regional seafloor imagery datasets from 19 Antarctic research cruises between 1985 and 2019. AS-AID comprises of 3,599 georeferenced downward facing seafloor images that have been labelled with a total of 615,051 expert annotations. Annotations are based on the CATAMI (Collaborative and Automated Tools for Analysis of Marine Imagery) classification scheme and have been reviewed by experts. In addition, because the pixel location of each annotation within each image is available, annotations can be viewed easily and customised to suit individual research priorities.

This dataset can be used to investigate species distributions, community patterns, it provides a reference to assess change through time, and can be used to train algorithms to automatically detect and annotate marine fauna.

Background & Summary

Marine imagery is a cost-effective and non-intrusive way to gather information about the seafloor. Extracting sound and detailed biological information from these images, however, is a key bottleneck because image annotation is time-consuming and often expensive. Perhaps because these data are difficult to produce, they have historically rarely been shared on public repositories. There has been a recent societal and scientific shift towards sharing data, greater transparency and reproducibility of data (FAIR data principles)[1,2], and increasing capabilities of machine learning methods that require large numbers of annotated images has encouraged more scientists and research institutions to make datasets of expert annotated marine images publicly available (e.g., FathomNet[3]). Nonetheless, for most regions around the world, including the Southern Ocean, expert annotated seafloor images are rarely publicly available.

The sparsity of biological data impedes efforts to benchmark the status of marine ecosystems, and therefore also prevents monitoring of ecological change. This is particularly evident for the highly biodiverse Antarctic seafloor[4-6], which still lacks a comprehensive analysis of the distribution of its biodiversity on a continental scale. While Southern Ocean ecosystems are predicted to change significantly as global temperatures change[7-11], the information underlying this knowledge stems from either single regions[8,10,12] or global models[11], from single taxonomic groups[9] or from datasets that compile opportunistically collected fauna[7]. A database of the distribution and abundances of a broad range of organisms covering the entire Antarctic continental shelf is invaluable to help identifying key conservation areas and managing these unique ecosystems.

In this article, we introduce a database of expert annotated seafloor imagery from the Antarctic continental shelf and slope, the Antarctic Seafloor Annotated Imagery Database (AS-AID, Fig. 1). AS-AID comprises of 3,599 annotated images and 615,051 annotations of a total of 180 faunal classifications. Images have been collected from scientific surveys between 1985 and 2019 for which data was either publicly available or could be sourced by directly contacting data curators. This represents the first international effort to collate images from different regions of the Southern Ocean at a pan-Antarctic scale into a single database and annotate these images using a single classification scheme.

Repeatability is key to making data useful beyond the scope of a single project, which is why annotations in AS-AID are based on an internationally recognised standard for labelling marine images (CATAMI[13]), and all annotations have exact reference points (x and y coordinates) to their respective images. Users can view AS-AID annotations on Squidle+ (www.squidle.org) and review all annotations for a single label and customise annotation labels to suit their own research objective. AS-AID can be used to train deep learning algorithms for automated classification of marine fauna. Combining AS-AID with environmental data and statistical models can allow insights into the species distributions and ecological processes that shape Antarctic seafloor ecosystems[10,14-16].

Methods

The Antarctic Seafloor Annotated Image Database (AS-AID) includes downward facing images of the seafloor at a depth range between 100 – 3,000 m from Antarctic scientific surveys for which we were able to obtain access to both images and detailed metadata records. We restricted the database to images deeper than 100 m because at these depths environmental conditions are more stable (e.g., relatively less impact on the fauna from iceberg scour) and more comparable between regions, which suits the continental scale biodiversity analysis that this database has primarily been developed for. AS-AID focusses on downward facing seafloor imagery rather than oblique facing imagery (where cameras can be mounted at different

angles) because initial work indicated that downward facing images are more comparable between different surveys, with greater uniformity of illumination and resolution. Additionally, the area of seabed imaged with downward facing cameras can be more accurately estimated than from data collected with obliquely mounted cameras. In total, we identified 19 surveys from 2019 and prior that collected downward facing images of suitable quality with accessible metadata records also available (Fig. 1, 2 & Table 1). The only surveys that we identified as potentially suitable, but for which we could not access all relevant data were LMG1703 and ANA08D, which are therefore excluded from AS-AID at time of writing.

Sampling locations and image collection

AS-AID images were collected during 19 scientific cruises between 1985 and 2019 (Tables 1-4), using a range of different camera platforms (Table 4). Figure 2 shows an overview of sampling locations for all image transects included in AS-AID. Images and metadata were downloaded and stored on external hard-drives, and image filenames were standardised for processing.

Image pre-processing

We inspected all 72,560 images visually and removed images/transects from AS-AID that were out of focus, too far or too close to the seafloor to allow identification of animals, had poor lighting, or had unreliable coordinates (Tables 2 & 3). We then corrected images for tone, colour and contrast using the automated batch function in Adobe Photoshop Version 21 (all surveys apart from PS96, see Table 4). Further, because the corners of images from surveys PS81, PS96, PS118 and AA2011 are out of focus we cropped these images to either 80% or 90% of their original size (Table 4).

Image sub-setting

Annotating all images from all surveys is neither feasible nor desirable due to the large differences in sampling effort among surveys (i.e., the density of images along the transect lines is highly variable). We therefore created a subset of the full dataset by selecting a spatially balanced random subset of the images using their coordinates and the distance along the transect line with the R package MBHdesign version 2.2.2. Where only one coordinate for the transect was available, or where recorded end-coordinates of the transects did not match with expected values given the duration of the deployment and the normal ship-speed during deployment (max 1 kn), we randomly selected 5 of the first 25 images in those transects and assigned these images to the starting position of the transect (Table 2).

The primary image subset we created contains approximately 1 image for every 100 m of survey transect (3,599 images, Table 3). Further, we created a smaller secondary subset of images for a more detailed exhaustive search of each image, containing approximately 1 image for every 200 m of survey transect (1912 images, Table 3).

Image annotation

We used both a point-grid and an exhaustive search for annotating the images. Point-grid annotation allows estimating percentage cover of all substrates and organisms present in the image, but the number of points that make up the grid influence how well these percentages are estimated, and how well rare and/or small organisms are detected[17]. In contrast, labelling organisms individually in an exhaustive search reduces detectability issues for rare species. However, individual labels present a challenge for many colonial organisms because individual colony sizes can vary greatly and it is frequently ambiguous how many colonies there are. We chose to first label all images in the primary subset using a point-grid, and then label the secondary subset using an exhaustive search. In the exhaustive search we focussed only on two groups of organisms: Mobile species, and small or rare species comprising Vulnerable

Marine Ecosystem taxa (VME taxa as defined by the Commission for the Conservation of Antarctic Marine Living Resources, CCAMLR) that are not colonial.

For point-grid annotation we overlaid all images in the primary subset with a regular grid of 108 points (9x12) to closely match the 3x4 height to width ratio of the images. We chose 108 points as a good compromise between sampling effort and time constraints. For each point we only identified what substrate or organisms could be found within the single pixel in the centre of this point. While this required us to regularly zoom in and out of the image to determine the exact position of the point, it reduced potential bias towards labelling a point as an organism instead of substrate. We used CoralNet (coralnet.ucsd.edu/), which is a web-based toolkit with inbuilt deep-learning capabilities[18]. We manually annotated all images and used the inbuilt deep-learning algorithm as assistance to more quickly select the correct label.

For individually labelling organisms in the secondary subset via an exhaustive search, we used the online toolkit BIIGLE 2.0 (www.biigle.de/) and their "lawnmower mode" to systematically search through an image[19]. We identified each organism from the respective morphospecies list (Supplementary Classification Catalog) and drew a close-fitting circle around the organism if they were mainly (>50% of their body) within the image.

For all images in the primary subset, we used BIIGLE to annotate visible laser-points that allow calculating the imaged seabed area.

Image classification scheme

We classified all organisms based on CATAMI[13] (Collaborative and Annotation Tools for Analysis of Marine Imagery), and added labels to the classification tree when we could confidently distinguish organisms further into separate morphospecies.

CATAMI is a hierarchical classification scheme that uses taxonomy at the higher levels of the classification tree and morphology at the lower levels of the classification tree. This approach accounts for the fact that it is often difficult to identify seafloor fauna to a fine taxonomic resolution without physical specimens and that the morphology of seafloor fauna is often reflective of function. Where further discrimination is known or needed these can be added at the end of the respective CATAMI branch.

Our classification catalogue for both the point-scoring and the exhaustive search can be found in the supplementary information.

Quality control & expert review

All annotations went through rigorous quality control comprising three stages: Labelling review, internal QA/QC, and expert review.

Labelling review concerns the exhaustive search only: After all images of a survey had been fully annotated in BIIGLE, a second scorer systematically searched through all images again to label any organisms the first scorer had missed. The use of two scorers to search though each image gives us high confidence that few individuals have been missed.

Internal QAQC: We conducted a two-stage internal QAQC both individually for each survey during annotation and also for all surveys combined. In the internal QAQC we reviewed each label, corrected any misclassification and discussed with the team if necessary. For reviewing labels from the grid annotation, we cropped 50x50 pixel thumbnails around each grid-point and sorted the thumbnails by their label in a morphospecies library. For reviewing the individual annotations from the exhaustive search using BIIGLE we used BIIGLE's inbuilt "Largo" review function. After this first review, we then created a second library containing only thumbnails of morphospecies whose labels had changed during the review. This second review stage was to ensure all changes made to the original scoring were correct.

Expert review: In the third stage of our review process, we consulted with taxonomic experts to ensure the morphospecies grouping and labelling was consistent, and to help identify some of the unknown organisms encountered.

Data Records

Link 1: To be included here is a link to a persistent and publicly accessible repository that contains all images, metadata and annotations of AS-AID.

Link 2: To be included here is a link to the data on SQUIDLE+. SQUIDLE+ has a user friendly interface that can be used to interrogate the annotated images in more detail.

Technical Validation

Annotations in AS-AID have gone through a rigorous review process involving experienced scientists, trained students and taxonomists. During the annotation process, the annotation team (consisting of two scientists, one postgraduate student and one undergraduate student) met regularly (daily at the beginning, weekly at a later stage and every time a survey was fully annotated) to discuss uncertain annotations, ensure consistent annotation between the scorers and create new labels where necessary. All annotations were checked for correct labelling once a survey was fully annotated and the entire database was checked again after all surveys had been fully annotated. Further, expert taxonomists were consulted in particular for key taxa such as sponges, cnidarians and bryozoans. The largest uncertainties in this database were distinguishing whether an object was an organism or substrate, and whether an organism was dead or alive, and we created a number of categories to label these cases accordingly, such as e.g. “Unknown Biological Matrix” or “Unidentifiable”.

In AS-AID, we have opted to use experts to consecutively perform tasks on the same set of images. In particular, this meant one expert labelled the images while a second expert checked the annotations in the review process and flagged any misclassifications for discussion. While this process didn’t allow label accuracy to be quantified, it did allow organisms to be labelled more consistently throughout the entire database.

Researchers are invited to assess the validity of AS-AID for their own research purposes. Every annotation in AS-AID is linked to an image with a unique filename and contains the x- and y-coordinates where the organisms is found on that image as well as other metadata such as imaged area and imaging gear. Therefore, every annotation is fully traceable, allowing researchers to customise and assess the validity of AS-AID for their needs.

Acknowledgements

We thank the crews and onboard scientific parties of the expeditions which collected the data within this study for providing the images and metadata for analysis in this study, either directly or via online raw data archives. We thank Amy Leventer for helping with data acquisition and Fanny Hermand for her help during the early stages of this project. This research has been funded as part of ARC Discovery Project 190101858. C.G is supported by a CSIRO-UTAS Quantitative Marine Science Scholarship. NZ involvement is supported by Ministry of Business Innovation and Employment programme C01X1710, *Ross Sea Research and Monitoring Programme; is the world’s largest MPA effective?*

Author contributions

JJ: Conceptualization, Methodology, Software, Validation, Investigation, Data Curation, Writing - Original Draft, Writing – Review & Editing, Visualization, Supervision, Project Administration, Funding acquisition

VS: Methodology, Validation, Investigation, Data Curation, Writing – Review & Editing

CG: Methodology, Validation, Investigation, Data Curation, Writing – Review & Editing

TW: Methodology, Validation, Investigation, Writing – Review & Editing

NH: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project Administration, Funding acquisition

DKB: Investigation, Resources, Writing – Review & Editing

DAB: Investigation, Resources, Writing – Review & Editing

JG: Investigation, Resources, Writing – Review & Editing

NB: Investigation, Writing – Review & Editing

RD: Investigation, Writing – Review & Editing

ME: Investigation, Writing – Review & Editing

ALP: Investigation, Writing – Review & Editing

HG: Resources, Writing – Review & Editing

KL: Resources, Writing – Review & Editing

AP: Resources, Writing – Review & Editing

DP: Investigation, Resources, Writing – Review & Editing

CS: Resources, Writing – Review & Editing

AZ: Resources, Writing – Review & Editing

CJ: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project Administration, Funding acquisition

Competing interests

The authors declare no competing financial interests.

Figures

Figure 1:

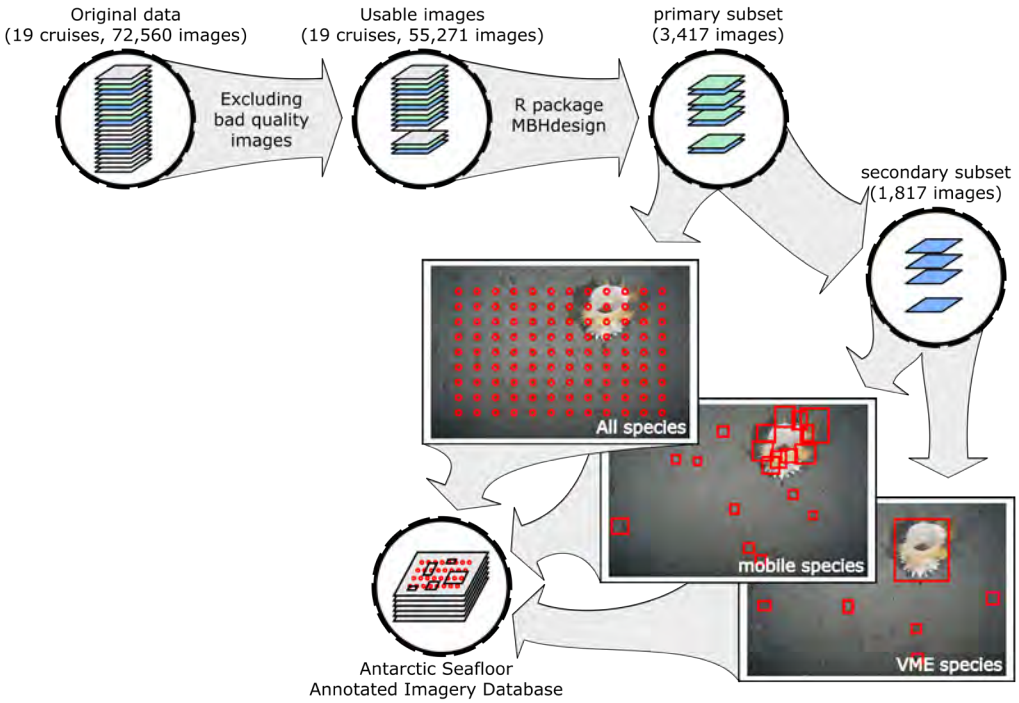

Figure 2:

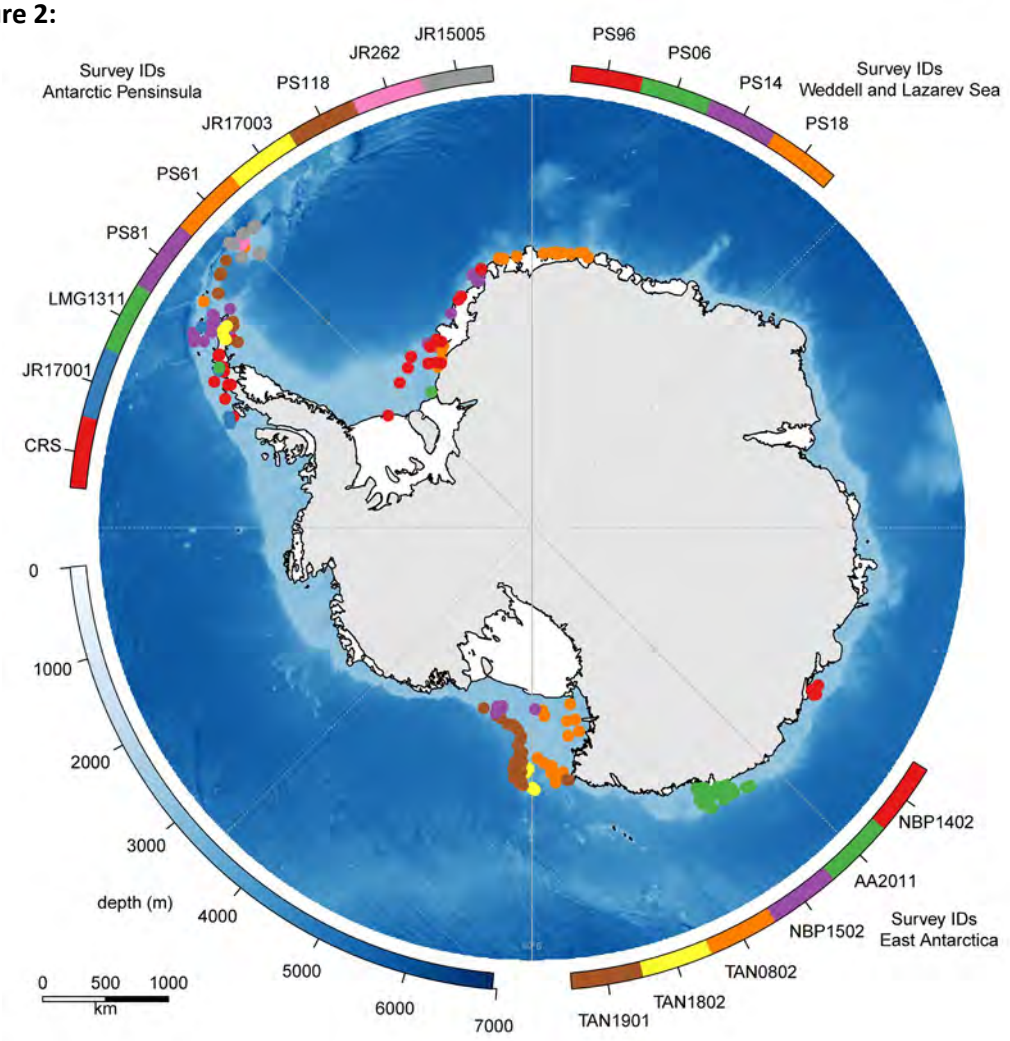

Figure 3:

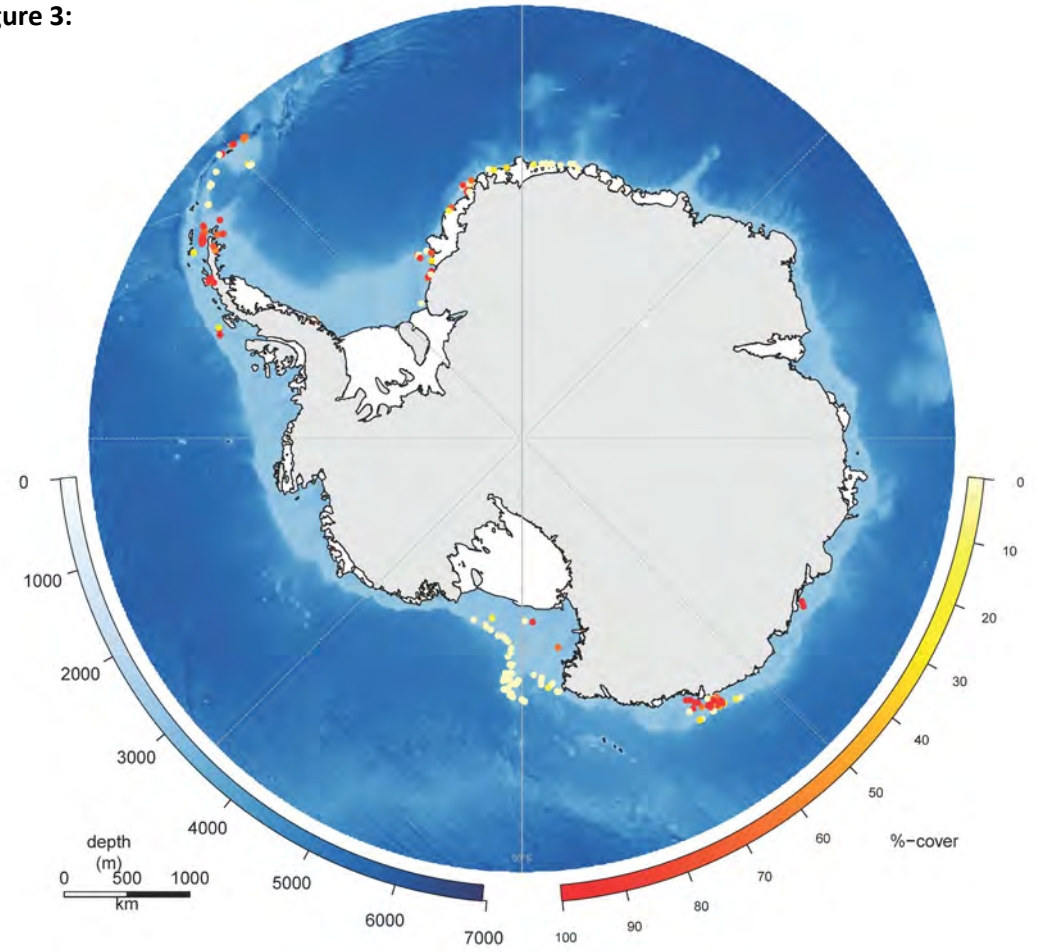

Figure 4:

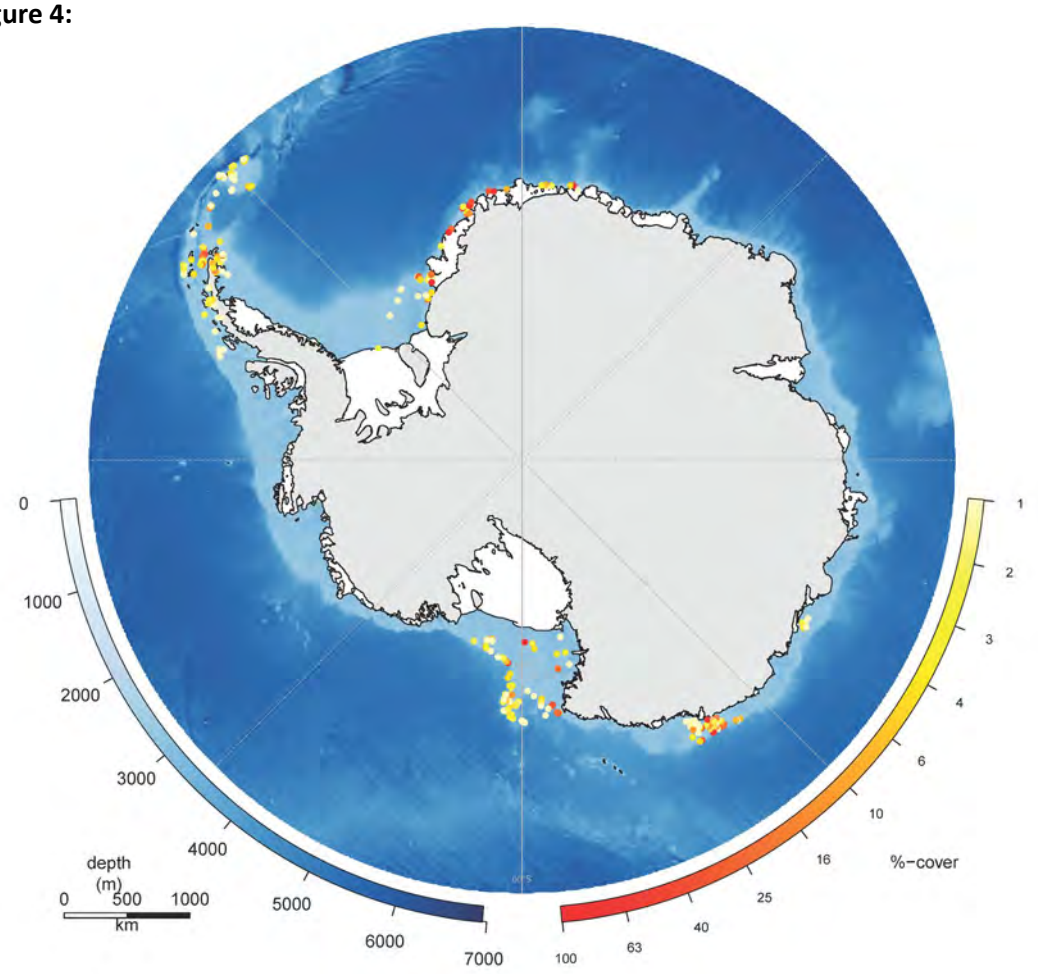

Figure Legends

Figure 1: Simple overview of the steps involved in building AS-AID, including the methodology used to generate the database. Seabed photo sourced from pangea.de

Figure 2: Overview of the circum-Antarctic distribution of sampling locations. Details about each survey can be found in tables 1-4.

Figure 3: Percentage cover of fine substrate across all sampling locations

Figure 4: Combined percentage cover of all ascidians, bryozoans, cnidarians, hydrocorals and sponges at each sampling location.

Tables

Table 1: Survey details, including links and references to data repositories

Survey ID	Year	Region	Contact	Reference/repository for data	Dataset link (if available)	License (if available)
PS06 (ANT-III/3)	1984/85	Weddell Sea	J. Gutt	Gutt [20],Gutt [21]	www.pangaea.de/?q=ps06+benthos&f.author%5B%5D=Gutt%2C+Julian&f.pubyear%5B%5D=1985	CC-BY-3.0
PS14 (ANT-VII/4)	1989	Weddell Sea	J. Gutt	Gutt [22],Gutt [23],Gutt [24],Gutt [25],Gutt [26],Gutt [27],Gutt [28],Gutt [29],Gutt [30],Gutt [31],Gutt [32],Gutt [33],Gutt [34],Gutt [35],Gutt [36],Gutt [37],Gutt [38],Gutt [39]	www.pangaea.de/?q=ps14+benthos&f.author%5B%5D=Gutt%2C+Julian&f.pubyear%5B%5D=2010&f.pubyear%5B%5D=1989	CC-BY-3.0
PS18 (ANT-IX/3)	1990/91	Lazarev Sea	J. Gutt	Gutt [40],Gutt [41],Gutt [42],Gutt [43],Gutt [44],Gutt [45],Gutt [46],Gutt [47],Gutt [48],Gutt [49],Gutt [50],Gutt [51],Gutt [52],Gutt [53],Gutt [54],Gutt [55],Gutt [56],Gutt [57],Gutt [58],Gutt [59],Gutt [60],Gutt [61]	www.pangaea.de/?q=ps18+benthos&f.author%5B%5D=Gutt%2C+Julian	CC-BY-3.0
PS61 (ANT-XIX/5)	2002	Antarctic Peninsula	J. Gutt	Arntz and Gutt [62],Arntz and Gutt [63]	doi.pangaea.de/10.1594/PANGAEA.220747 doi.pangaea.de/10.1594/PANGAEA.220746	CC-BY-3.0
PS81 (ANT-XXIX/3)	2013	Antarctic Peninsula	D. Piepenburg	Piepenburg, et al. [64]	doi.pangaea.de/10.1594/PANGAEA.872719	CC-BY-3.0
PS96 (ANT-XXXI/2)	2015/16	Weddell Sea	D. Piepenburg	Piepenburg [65]	doi.pangaea.de/10.1594/PANGAEA.862097	CC-BY-3.0
PS118	2019	Antarctic Peninsula	A. Purser	Purser, et al. [66]	doi.pangaea.de/10.1594/PANGAEA.911904	CC-BY-4.0
TAN0802	2008	Ross Sea	D. Bowden		https://atlas.niwa.co.nz/	CC-BY-3.0
TAN1802	2018	Ross Sea	D. Bowden		https://atlas.niwa.co.nz/	CC-BY-3.0
TAN1901	2019	Ross Sea	D. Bowden		https://atlas.niwa.co.nz/	CC-BY-3.0
CRS	various cruises/ years	Antarctic Peninsula	C. Smith A. Ziegler	Grange and Smith [67]	Dataset link to be made available	
NBP1402	2014	East Antarctica	A. Post	Leventer, et al. [68],Post, et al. [69]	https://www.usap-dc.org/view/dataset/601310	CC BY-NC 4.0
AA2011	2011	East Antarctica	J. Smith	Smith [70]	data.aad.gov.au/metadata/records/VMS_Benthic_Photography	CC-BY-4.0
NBP1502	2015	Ross Sea	P. Bart	Bart [71]	https://doi.org/10.15784/601182	CC BY-NC 4.0
LMG1311	2013	Antarctic Peninsula	A. Leventer	Domack [72]	https://doi.org/10.15784/601311	CC BY-NC 4.0
JR262	2011	South Orkney	D. Barnes	Barnes and Sands [73],Barnes, et al. [74]		
JR15005	2015	South Orkney	H. Griffiths	Brasier, et al. [75]		
JR17001	2017	Antarctic Peninsula	D. Barnes	Barnes, et al. [76],Zwerschke, et al. [77]		
JR17003	2018	Antarctic Peninsula	K. Linse	Linse, et al. [78]	doi.org/10.5285/48dcef16-6719-45e5-a335-3a97f099e451	http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/

Table 2: Data characteristics for each cruise and comments on which data were excluded. Metadata quality: * single lon/lat position for each transect; ** lon/lat available for start/end locations of each transect Locations of individual images interpolated from start/end locations; *** lon/lat available for each image

Survey ID	Metadata quality	Transects included	Comments
PS06	*	2/10	Excluded transect 288: located near/under iceshelf (no available environmental data) Excluded transects 301-303, 307, 309-311: Blurry images
PS14	**	18/24	Image quality low Excluded transects 248, 250, 259, 261: Blurry/discoloured images Excluded transect 274: Coordinates unreliable (transect on land) Excluded transect 312: Located in the South Atlantic
PS18	**	22/25	Excluded transects 162-1, 173-2 & 207-1: located near/under iceshelf (no available environmental data)
PS61	**	2/8	Only included transects 235-1 & 249-1: all other transects are on the Scotia Arc which is not part of this database
PS81	***	29/30	Excluded transect 159: Bad visibility making it difficult to identify organisms
PS96	***	13/13	-
PS118	***	10/11	Excluded transect 38: Bad visibility making it difficult to identify organisms Transects 39 & 69: Reduced transect starts and ends to form straight transect lines
TAN0802	***	17/53	Excluded transects 181-310: located on seamounts off the Antarctic continental shelf
TAN1802	***	31/31	Planned to exclude transects 160-213 which are located on seamounts off the Antarctic continental shelf. However, accidently annotated images from these 16 transects anyway.
TAN1901	***	34/34	Planned to exclude transect 209 which is located on a seamount off the Antarctic continental shelf. However, accidently annotated images from this transects anyway.
CRS	**	32/34	Excluded transect 1103: Bad illumination/image quality Excluded transect 1325: Uncertain reliability of coordinates (transect very long with few images only)
NBP1402	**	10/10	-
AA2011	***	71/86	Excluded transects CTD53, CTD55, CTD57-58, CTD61-64, CTD85, CTD90, CTD93, CTD98, CTD107, CTD115 & CTD138: no or only blurry/unusable images
NBP1502	**	5/7	Excluded transects KC16 & KC17: no metadata available/accessible
LMG1311	**	2/2	-
JR262	*	4/23	Image quality very low 4 transects on the South Orkney shelf included, all other transects are off the Antarctic continental shelf/slope
JR15005	**	21/21	-
JR17001	***	7/7	-
JR17003	***	12/12	-

Table 3: Overview of the number of images per survey and how subsets were created. “Total” is the number of images downloaded from the repository; “Usable” is the number of images that show the seafloor, are in focus and are neither blurry nor too bright/dark; “Included” is the number of images added to our database, that is images that are from transects that we were interested in analysing; and “Subset 1” and “Subset 2” are the number of images that were selected for annotation for the point-scoring method and the exhaustive search respectively.

Survey ID	Number of images					Image subsetting
	Total	Usable	Included	Subset 1	Subset 2	
PS06	500	31	16	9	5	Transect 289: selected 5 random images Transect 292: selected all good quality images (4)
PS14	1523	1220	1098	119	65	Spatially-balanced-random However: Start/end coordinates not trustworthy (ship speed would be >1kt) for transects 245, 250, 256, 259, 260, 261, 275, 278-3, 280 & 293: Selected 5 random images out of the first 25 images from these transects, using starting GPS coordinates as image position.
PS18	1705	1654	1456	103	59	Spatially-balanced-random However: Start/end coordinates not trustworthy (ship speed would be >1kt) for transects 171-1, 173-1, 206-2, 211-2 & 222-1: Selected 5 random images out of the first 25 images from these transects, using starting GPS coordinates as image position.
PS61	139	129	129	12	7	Spatially-balanced-random
PS81	14543	14516	14021	1041	528	Spatially-balanced-random
PS96	2736	2727	2727	217	112	Spatially-balanced-random
PS118	12629	6057	4691	260	133	Spatially-balanced-random Transect 11-2: GPS-coordinates wiggle a lot between images -> subset based on distance from start rather than distance along transect Transect 39-1: removed the first 800 and the last 487 images to result in a straight lined transect Transect 69-1: removed the first 29 and the last 50 images to result in a straight lined transect
TAN0802	13046	11403	3310	173	91	Spatially-balanced-random
TAN1802	7470	5616	5616	336	175	Spatially-balanced-random
TAN1901	8301	5878	5878	373	195	Spatially-balanced-random
CRS	2396	2032	1982	320	169	Spatially-balanced-random
NBP1402	761	692	453	78	43	Spatially-balanced-random
AA2011	1853	435	435	71	71	Selected 1 random image per transect due to low number of usable images
NBP1502	2869	1719	1719	245	124	Spatially-balanced-random
LMG1311	263	222	222	60	31	Spatially-balanced-random
JR262	61	31	31	17	10	Selected 5 random images per transect
JR15005	773	505	505	88	53	Spatially-balanced-random
JR17001	142	120	120	14	8	Spatially-balanced-random
JR17003	850	279	279	63	33	Spatially-balanced-random
TOTAL	72560	55271	42090	3599	1912	

Table 4: Details on image collection, image characteristics and annotation

Survey ID	Camera-platform	Depth-range (IBCSO V2) for images		Distance from seafloor	Size of imaged seabed area	Laser-points	Image crop	Colour-corrected	Annotations		
		included	annotated						Point-score all species	Exhaustive search mobile species	Exhaustive search VME species
PS06	Photo sledge (FTS)	349 – 1115	same	Constant (weight)	Unknown	-	No	Yes	972	241	10
PS14	Photo sledge (FTS)	148 – 1347	same	Constant (weight)	0.56m^2	-	No	Yes	12852	1110	119
PS18	Photo sledge (FTS)	144 – 956	176 – 956	Constant (weight)	~0.9m^2 (but cruise report says 1m^2…)	-	No	Yes	11124	1889	180
PS61	Photo sledge (FTS)	330 – 348	same	Constant (weight)	~1m^2	-	No	Yes	1296	218	4
PS81	OFOS	52 – 783	same	Variable	Calculate	50cm	80%	Yes	112428	83888	15249
PS96	OFOS	198 – 731	same	Variable	Provided + calculated	50cm	80%	No	23436	7602	3852
PS118	OFOBS	82 – 2169	same	Variable	Calculate	50cm	80%	Yes	28080	24900	6692
TAN0802	DTIS	287 – 2258	same	Variable	Calculate	20cm	No	Yes	18684	6237	872
TAN1802	DTIS	664 – 1213	same	Variable	Calculate	20cm	No	Yes	36288	12279	5964
TAN1901	DTIS	447 – 1373	same	Variable	Calculate	20cm	No	Yes	40284	16837	9523
CRS	YOYO	239 – 693	same	Constant (weight, 2.5m)	~3m^2 + calculated	10cm	No	Yes	34560	5850	573
NBP1402	YOYO	281 – 899	same	Constant (weight, 2.5m)	~4.8m^2 + calculated	10cm	No	Yes	8424	2474	342
AA2011	CTD	198 – 1453	same	Variable	Calculate	50cm	90%	Yes	7668	5322	543
NBP1502	YOYO	175 – 597	same	Calculate	Calculate	10cm	No	Yes	26460	6242	1281
LMG1311	YOYO	349 – 654	same	Calculate	Calculate	10cm	No	Yes	6480	1720	308
JR262	SUCS	219 – 280	same	Constant (frame)	0.51m^2	-	No	Yes	1836	112	9
JR15005	SUCS	473 – 1017	same	Constant (frame)	0.51m^2	-	No	Yes	9504	875	112
JR17001	SUCS	67 – 507	same	Constant (frame)	0.51m^2	-	No	Yes	1512	74	21
JR17003	SUCS	184 – 865	same	Constant (frame)	0.51m^2	-	No	Yes	6804	2702	133
TOTAL		67 – 2258							388692	180572	45787

Supplementary information

The size of the classification catalog files is larger than the maximum allowed size for files at BioRxiv. The classification catalog files are available upon request from the corresponding author and will be made publicly available with the peer-reviewed version of this paper.

References

1 Schoening, T. *et al.* Making marine image data FAIR. *Scientific Data* **9**, doi:10.1038/s41597-022-01491-3 (2022).

2 Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* **3** (2016).

3 Katija, K. *et al.* FathomNet: An Open, Underwater Image Repository for Automated Detection and Classification of Midwater and Benthic Objects. *Mar. Technol. Soc. J.* **55**, 136-137, doi:10.4031/mtsj.55.3.20 (2021).

4 Griffiths, H. J., Barnes, D. K. A. & Linse, K. Towards a generalized biogeography of the Southern Ocean benthos. *J. Biogeogr.* **36**, 162-177, doi:10.1111/j.1365-2699.2008.01979.x (2009).

5 De Broyer, C. *et al. Biogeographic Atlas of the Southern Ocean*. (Scientific Committee on Antarctic Research, 2014).

6 Chown, S. L. *et al.* The changing form of Antarctic biodiversity. *Nature* **522**, 431-438, doi:10.1038/nature14505 (2015).

7 Griffiths, H. J., Meijers, A. J. S. & Bracegirdle, T. J. More losers than winners in a century of future Southern Ocean seafloor warming. *Nature Climate Change* **7**, 749, doi:10.1038/Nclimate3377 (2017).

8 Pineda-Metz, S. E. A., Gerdes, D. & Richter, C. Benthic fauna declined on a whitening Antarctic continental shelf. *Nat. Commun.* **11**, doi:10.1038/s41467-020-16093-z (2020).

9 Fabri-Ruiz, S. *et al.* Benthic ecoregionalization based on echinoid fauna of the Southern Ocean supports current proposals of Antarctic Marine Protected Areas under IPCC scenarios of climate change. *Global Change Biol.* **26**, 2161-2180, doi:10.1111/gcb.14988 (2020).

10 Jansen, J. *et al.* Mapping Antarctic Suspension Feeder Abundances and Seafloor Food-Availability, and Modeling Their Change After a Major Glacier Calving. *Frontiers in Ecology and Evolution* **6**, 94, doi:10.3389/fevo.2018.00094 (2018).

11 Jones, D. O. *et al.* Global reductions in seafloor biomass in response to climate change. *Global Change Biol.* **20**, 1861-1872, doi:10.1111/gcb.12480 (2014).

12 Marini, S. *et al.* Long-term High Resolution Image Dataset of Antarctic Coastal Benthic Fauna. *Scientific Data* **9**, doi:10.1038/s41597-022-01865-7 (2022).

13 Althaus, F. *et al.* A standardised vocabulary for identifying benthic biota and substrata from underwater imagery: the CATAMI classification scheme. *PloS one* **10**, e0141039 (2015).

14 Jansen, J., Hill, N. A., Dunstan, P. K., Eléaume, M. P. & Johnson, C. R. Taxonomic Resolution, Functional Traits, and the Influence of Species Groupings on Mapping Antarctic Seafloor Biodiversity. *Frontiers in Ecology and Evolution* **6**, 81, doi:10.3389/fevo.2018.00081 (2018).

15 Jansen, J. *et al.* Abundance and richness of key Antarctic seafloor fauna correlates with modelled food availability. *Nature Ecology and Evolution* **2**, 71-80, doi:10.1038/s41559-017-0392-3 (2018).

16 Jansen, J. *et al.* Integrated assessment of the spatial distribution and structural dynamics of deep benthic marine communities. *Ecol. Appl.* **30**, e02065, doi:10.1002/eap.2065 (2020).

17 Perkins, N. R., Foster, S. D., Hill, N. A. & Barrett, N. S. Image subsampling and point scoring approaches for large-scale marine benthic monitoring programs. *Estuar. Coast. Shelf Sci.* **176**, 36-46, doi:10.1016/j.ecss.2016.04.005 (2016).

18 Williams, I. D. *et al.* Leveraging automated image analysis tools to transform our capacity to assess status and trends on coral reefs. *Frontiers in Marine Science* **6**, 222 (2019).

19 Langenkämper, D., Zurowietz, M., Schoening, T. & Nattkemper, T. W. BIIGLE 2.0 - Browsing and Annotating Large Marine Image Collections. *Frontiers in Marine Science* **4**, doi:10.3389/fmars.2017.00083 (2017).

20 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS06/289. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.713324 (1985).

21 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS06/292. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.713325 (1985).

22 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/245. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691559 (1989).

23 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/246. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691560 (1989).

24 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/256. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691562 (1989).

25 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/260. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691564 (1989).

26 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/270. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691566 (1989).

27 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/275. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691568 (1989).

28 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/276. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691569 (1989).

29 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/277. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691570 (1989).

30 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/278-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691571 (1989).

31 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/278-3. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691572 (1989).

32 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/280. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691573 (1989).

33 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/285. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691574 (1989).

34 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/293. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691575 (1989).

35 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/294. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691576 (1989).

36 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/304. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691577 (1989).

37 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/305. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691578 (1989).

38 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/306. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691579 (1989).

39 Gutt, J. Sea-bed photographs (benthos) from the Weddell Sea (Antarctica) along profile PS14/307. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.691580 (1989).

40 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/126-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667008 (2007).

41 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/126-6. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667009 (2007).

42 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/129-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667010 (2007).

43 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/131-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667011 (2007).

44 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/134-1. *Alfred Wegener Institute, Helmholtz Centre for*

Polar and Marine Research, Bremerhaven, PANGAEA, doi:10.1594/PANGAEA.667012 (2007).

45 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/135-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667013 (2007).

46 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/136-4. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667014 (2007).

47 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/160-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667015 (2007).

48 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/165-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667017 (2007).

49 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/169-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667018 (2007).

50 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/171-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667019 (2007).

51 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/174-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667021 (2007).

52 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/175-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667022 (2007).

53 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/179-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667023 (2007).

54 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/180-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667024 (2007).

55 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/182-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667025 (2007).

56 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/189-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667026 (2007).

57 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/206-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667027 (2007).

58 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/211-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667029 (2007).

59 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/212-2. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667030 (2007).

60 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/220-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667031 (2007).

61 Gutt, J. Sea-bed photographs (benthos) from the Weddell and Lazarev Seas (Antarctica) along profile PS18/222-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.667032 (2007).

62 Arntz, W. E. & Gutt, J. Sea-bed photographs (benthos) from the Scotia Arc (Antarctica) along profile PS61/235-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.220746 (2004).

63 Arntz, W. E. & Gutt, J. Sea-bed photographs (benthos) from the Scotia Arc (Antarctica) along profile PS61/249-1. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.220747 (2004).

64 Piepenburg, D., Segelken-Voigt, A. & Gutt, J. Seabed photographs taken along OFOS profiles during POLARSTERN cruise PS81 (ANT-XXIX/3). *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA,*, doi:10.1594/PANGAEA.872719 (2013).

65 Piepenburg, D. Seabed photographs taken along OFOS profiles during POLARSTERN cruise PS96 (ANT-XXXI/2 FROSN). *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA*, doi:10.1594/PANGAEA.862097 (2016).

66 Purser, A., Hehemann, L., Dreutter, S., Dorschel, B. & Nordhausen, A. OFOBS Seafloor images from the Antarctic Peninsula and Powell Basin, collected during RV POLARSTERN cruise PS118. *Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA,*, doi:10.1594/PANGAEA.911904 (2020).

67 Grange, L. J. & Smith, C. R. Megafaunal communities in rapidly warming fjords along the West Antarctic Peninsula: hotspots of abundance and beta diversity. *PloS one* **8**, e77917, doi:10.1371/journal.pone.0077917 (2013).

534 68 Leventer, A. *et al.* Bottom photos from the Southern Ocean acquired during R/V
535 Nathaniel B. Palmer expedition NBP1402. *U.S. Antarctic Program (USAP) Data Center*,
536 doi:doi.org/10.1594/IEDA/324208 (2020).

537 69 Post, A. L. *et al.* Environmental drivers of benthic communities and habitat
538 heterogeneity on an East Antarctic shelf. *Antarct. Sci.* **29**, 17-32,
539 doi:10.1017/s0954102016000468 (2017).

540 70 Smith, J. High resolution still photographs of the seafloor across the Mertz Glacier
541 Region, Ver. 1. *Australian Antarctic Data Centre*, doi:doi:10.4225/15/59acda196ccfb
542 (2017).

543 71 Bart, P. NBP1502 YoYo camera benthic images from Ross Sea. *U.S. Antarctic Program*
544 *(USAP) Data Center*, doi:doi.org/10.15784/601182 (2019).

545 72 Domack, E. W. Processed Camera Images acquired during the Laurence M. Gould
546 expedition LMG1311. *U.S. Antarctic Program (USAP) Data Center*,
547 doi:doi.org/10.15784/601311 (2020).

548 73 Barnes, D. K. A. & Sands, C. J. Functional group diversity is key to Southern Ocean
549 benthic carbon pathways. *PloS one* **12**, e0179735, doi:10.1371/journal.pone.0179735
550 (2017).

551 74 Barnes, D. K. A. *et al.* Biodiversity signature of the Last Glacial Maximum at South
552 Georgia, Southern Ocean. *J. Biogeogr.* **43**, 2391-2399, doi:10.1111/jbi.12855 (2016).

553 75 Brasier, M. J. *et al.* Benthic biodiversity in the South Orkney Islands Southern Shelf
554 Marine Protected Area. *Biodiversity*, 1-15, doi:10.1080/14888386.2018.1468821
555 (2018).

556 76 Barnes, D. K. A. *et al.* Blue carbon gains from glacial retreat along Antarctic fjords:
557 What should we expect? *Global Change Biol.* **26**, 2750-2755, doi:10.1111/gcb.15055
558 (2020).

559 77 Zwerschke, N. *et al.* Quantification of blue carbon pathways contributing to negative
560 feedback on climate change following glacier retreat in West Antarctic fjords. *Global*
561 *Change Biol.* **28**, 8-20, doi:10.1111/gcb.15898 (2022).

562 78 Linse, K. *et al.* Benthic seafloor images from Prince Gustav Channel and Duse Bay,
563 Eastern Antarctic Peninsula, March 2018. *Polar Data Centre, Natural Environment*
564 *Research Council, UK Research & Innovation*, doi:doi.org/10.5285/48dcef16-6719-
565 45e5-a335-3a97f099e451 (2020).

566

567