An integrated data compilation of data sources for the development of a marine protected area in the Weddell Sea

Katharina Teschke1,2, Hendrik Pehlke1,2, Volker Siegel1, Horst Bornemann1, Rainer Knust1, Thomas Brey1,2,4

1 Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar und Meeresforschung, Am Handelshafen 12, 27570 Bremerhaven, Germany
2 Helmholtz Institute for Functional Marine Biodiversity at the University Oldenburg (HIFMB), Ammerländer Heerstraße 231, 26129 Oldenburg, Germany
3 Thünen Institute of Sea Fisheries, Herwigstraße 31, 27572 Bremerhaven, Germany
4 University Bremen, Bibliothekstraße 1, 28359 Bremen, Germany

Correspondence to: Katharina Teschke (Katharina.Teschke@awi.de)

Abstract. The Southern Ocean may contribute a considerable part to the proposed global network of Marine Protected Areas (MPAs) that should cover about 10% of the world oceans in 2020. In the Antarctic, the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR) is responsible for this task, and currently Germany leads a corresponding scientific evaluation of the wider Weddell Sea region. Compared to other marine regions within the Southern Ocean, the Weddell Sea is exceptionally well investigated. A tremendous amount of data and information has been produced over the last four decades. Here, we give a systematic overview compilation of these all data sources collected that were acquired in the context of the Weddell Sea MPA planning process. The data compilation of data sources comprises data produced by scientists / institutions from more than twenty countries and were either available within our institutes, provided by our collaborators, downloaded via data portals, or transcribed from the literature. It is the first compilation data compilation for this area that includes abiotic data, such as bathymetry and sea ice, and ecological data from zooplankton, zoobenthos, fish, birds and marine mammals. The final All data layer products based on this huge data compilation of environmental and ecological data, including metadata description, are available from the data publisher PANGAEA via the five six persistent identifiers at https://doi.org/10.1594/PANGAEA.899595 (Pehlke and Teschke, 2019), https://doi.org/10.1594/PANGAEA.899667 (Teschke et al., 2019a), https://doi.org/10.1594/PANGAEA.899645 (Teschke et al., 2019b), https://doi.org/10.1594/PANGAEA.899591 (Teschke et al., 2019c), https://doi.org/10.1594/PANGAEA.899520 (Pehlke et al., 2019a), and https://doi.org/10.1594/PANGAEA.899595 (Pehlke and Teschke, 2019b), https://doi.org/10.1594/PANGAEA.899619 (Pehlke et al., 2019b), https://doi.org/10.1594/PANGAEA.899645 (Teschke et al., 2019b), and https://doi.org/10.1594/PANGAEA.899667 (Teschke et al., 2019c). This data compilation of data sources with the final data layer products will serve future research and monitoring well beyond the current MPA development process.

1 Introduction

Marine Protected Areas (MPAs) have experienced a significant increase in number and coverage at a global scale during recent decades (e.g. Mora and Sale, 2011; McDermott et al., 2018; UNEP-WCMC and IUCN,
The number of MPAs has increased almost 1.5 times since the 1990s and the total area protected is currently almost 30 million km². At the United Nations World Summit on Sustainable Development in 2002 the international community of states reached an agreement about the establishment of a representative network of MPAs for the purposes of long-term conservation of marine biodiversity by 2012 (A/CONF.199/20, 2002). The adopted strategic plan for biodiversity 2011-2020 of the Convention on Biological Diversity aims at the conservation of at least 10% of the coastal and offshore marine areas by 2020 based on a MPA network (CBD, 2010). The Southern Ocean may contribute a considerable proportion of this MPA network due to its size, and the uniqueness of the Antarctic environment renders its conservation the more urgent.

The Weddell Sea represents the southerly part of the Atlantic Sector of the Southern Ocean. About one quarter of the Weddell Sea’s entire marine area covers the continental shelf along the eastern contour of the Antarctic Peninsula and the Antarctic continent up to 20°E as a non topographic delineation. The Weddell Sea is deserving protection in multiple respects. On the one hand, all arguments for the conservation of the Southern Ocean hold true for the Weddell Sea, too: An extreme environment mostly dominated by the seasonal dynamic of the sea ice with an excellent adapted biota. The biodiversity is - particularly in the benthos - very high (e.g. Brey et al., 1994; Brandt et al., 2007), and there is a significant number of endemic species, i.e. unique to the Antarctic or even to the Weddell Sea (e.g. Arntz et al., 1994; Clarke and Johnston, 2003; Linse et al., 2006). Moreover, the Weddell Sea plays an important role for seabirds, penguins and marine mammals. Almost one third of the entire population of emperor penguins (Fretwell et al., 2012) and a major part of the circum-Antarctic population of crabeater seals (e.g. Bester and Odendaal, 2000; Southwell et al., 2012; Gurarie et al., 2016) apparently occurs in the Weddell Sea. Sponge associations which are comparable to tropical reef systems in terms of their structural and functional complexity occur along the eastern Weddell Sea shelf (Barthel and Gutt, 1992), and on the broad shelf in the southern Weddell Sea a special benthic community - adapted to very cold water temperatures - seems to resident (Teschke et al., 2016).

The Weddell Sea is - despite being one of the most remote and inaccessible places on earth - relatively well investigated compared to other Antarctic regions. Since approximately 30 years the Weddell Sea is the geographical focus area of the German Antarctic research. In addition, there are manifold research activities of other nations. Consequently, we were able to compile a tremendous amount of environmental and ecological data to support the development of a Weddell Sea MPA (hereafter: WSMPA) under the Commission for the Conservation of Antarctic Marine Living Resources (CCAMLR). Here we present a systematic overview of the integrated data compilation of environmental and ecological data sources collected for the development of a WSMPA and provide data layer products that are based on this data compilation.

2 Data description

2.1 Study site

The WSMPA Planning Area in which we acquired the environmental and ecological data is located between the Antarctic Peninsula and 20°E (Fig. 1). The northern border is at 64°S and the continental margin forms the southern border. This area is defined by CCAMLR's MPA Planning Domains in the CCAMLR Convention Area (SC-CAMLR-XXX, 2011) and by aiming at a bio-geographically homogeneous area, particularly on the shelf...
In addition to the WSMPA Planning Area (approx. 4.2 million km² in size) we compiled data for a 200 km wide buffer area near the Antarctic Peninsula, which is part of an MPA initiative led by Argentina and Chile (CCAMLR-XXXVII/31, 2018). This buffer zone is adjacent northerly to the northern border of the WSMPA Planning Area and has the eastern and western boundaries at 30°W and 60°W, respectively. Some data (e.g. seal tracking data), extend beyond the WSMPA Planning Area (plus buffer) and originate from adjacent regions of the Weddell Sea, such as the Bellinghausen Sea along the west side of the Antarctic Peninsula.

2.2 Data availability

All raw data sets of environmental and ecological parameters collected by the end of 2016 and further processed as part of the WSMPA planning process are systematically described and the primary reference is mentioned, such as the data portal on which the data are publically available, the institute/organisation on which the data can be requested on demand or the contact to the respective data provider (see Table 1 and 2; see all data records in Fig. 2 and Fig. S1). In addition, we offer data layer products that we developed on the basis of the raw data sets whose sources are described here. The methods used to process and analyse the data and to develop each data layer are described in detail in the Supplement. All raw data sets presented here were included in the WSMPA spatial planning analysis and the final data layer products with metadata description, including the description of analytical data processing, are freely available from the data publisher PANGAEA via the six persistent identifiers at https://doi.org/10.1594/PANGAEA.899595 (Pehlke and Teschke, 2019), https://doi.org/10.1594/PANGAEA.899667 (Teschke et al., 2019a), https://doi.org/10.1594/PANGAEA.899645 (Teschke et al., 2019b), https://doi.org/10.1594/PANGAEA.899591 (Teschke et al., 2019c), https://doi.org/10.1594/PANGAEA.899591 (Pehlke and Teschke, 2019), https://doi.org/10.1594/PANGAEA.899619 (Pehlke et al., 2019b), https://doi.org/10.1594/PANGAEA.899645 (Teschke et al., 2019b) and https://doi.org/10.1594/PANGAEA.899667 (Teschke et al., 2019a) (see Table 1 and 2). The data layers are available either as ArcMAP packages (as mxd file, containing a map document with all associated files) or as individual GIS files for those who use another GIS-software instead of the ESRI software (ArcMap). The shape and raster files, all with the same spheroid (WGS 1984) and projection (South Pole Lambert Azimuthal Equal Area, EPSG 102020), were processed in such a way that they can be easily used for the analysis of MPA scenarios or other geostatistical analyses in the Weddell Sea without direct access to the underlying raw data. For example, the shape and raster files could be stacked to identify hot- and coldspots of biodiversity, or certain layers could be used as explanatory variables in species distribution models.

2.3 Environmental data

2.3.1 IBCSO data

The bathymetric data used in the context of the WSMPA planning initiative originate from the first regional digital bathymetric model (DBM) established in the International Bathymetric Chart of the Southern Ocean.
(IBCSO) Version 1.0 programme and published by Arndt et al. (2013a, b) (data request: April 2013) (Table 1; Fig. 3a). This chart model is based upon bathymetric data of different origin, such as multi-beam and single beam data, digitized depths from nautical charts, predicted bathymetry, from many hydrographic offices, scientific institutions and data centres. The IBCSO Version 1.0 DBM has a horizontal resolution of 500 m x 500 m and a vertical resolution of 1 m based on a polar stereographic projection with true scale at 65° referenced to WGS84 ellipsoid (Arndt et al., 2013a, b).

2.3.2 AMSR-E sea ice maps

Daily high resolution sea ice maps of the Antarctic Ocean are provided by the PHAROS group (PHysical Analysis of RemOte Sensing images) at the Institute of Environmental Physics (IUP), University of Bremen, Germany. The sea ice raster maps, which were used in the context of the WSMPA planning initiative, are derived from satellite observations of daily sea ice concentration by the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-EOS) instrument on board the Aqua satellite. Daily AMSR-E sea ice concentration data (Jun 2002 - Oct 2011) were downloaded from IUP, University of Bremen (data request: 18-12-2013) (see Table 1; Fig. 3b). The ARTIST Sea Ice (ASI) concentration algorithm was used with a spatial resolution of 6.25 km x 6.25 km (Spreen et al., 2008) and a polar stereographic projection (EPSG: 3976).

2.3.3 FESOM data

Monthly mean values of seawater temperature, salinity and current velocity from 1990 to 2009 were derived from the Finite Element Sea Ice - Ocean Model (FESOM) (Table 1; Fig. 3c, d). The model run was initialised on January, 1st 1980 with hydrographic data from the Polar Science Center Hydrographic Climatology (Steele et al., 2001), and forced with NCEP daily atmospheric re-analysis data (Kalnay et al., 1996) for 1980 to 2009. For more information on FESOM and the atmospheric forcing data sets see e.g. Timmermann et al. (2009) and Haid and Timmermann (2013), respectively. The FESOM raster has a resolution of 0.18° (x) x 0.05° (y); in the vertical, two z-levels (i.e. sea surface and sea bottom) are used. The raster bases on WGS84 geographic coordinate system (EPSG: 4326).

IBCSO data, AMSR-E sea ice maps and FESOM data were used in a pelagic regionalisation analysis for the Weddell Sea. The respective data layer products are available at https://doi.org/10.1594/PANGAEA.899595 ("Pelagic regionalisation - clustering approach"). The clustering approach to classify different pelagic provinces is described in the Supplement. In addition, the data sets were used as environmental variables in various geostatistical approaches to develop spatial distribution maps for (i) adult Antarctic krill (AMSR-E), (ii) ice krill (IBCSO, FESOM), (iii) echinoderms (FESOM), (iv) demersal fish (IBCSO, FESOM), (v) Antarctic toothfish (IBCSO), (vi) Antarctic petrel (IBCSO, AMSR-E, FESOM) and (vii) emperor penguins (AMSR-E). The methods used to develop the different spatial distribution maps are described in the Supplement and the PANGAEA link to the respective data layer products (incl. file names) is given in the corresponding subsection under "2.4 Ecological data".
2.3.4 SeaWiFS data

Near-surface chlorophyll a concentration values stem from the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) measurements on board of the OrbView-2 (formerly SeaStar) spacecraft (Table 1). The monthly aggregated data (1997 to 2010) were downloaded via the NASA’s OceanColor website as level 3 standard mapped images (SMI) with a spatial resolution of 9 km x 9 km (data request: 09-09-2014).

2.3.5 WOA13 data

Data on dissolved oxygen, phosphate and nitrate were obtained from the World Ocean Atlas 2013 version 2 (WOA13 V2) (Garcia et al., 2014a, b) (Table 1). The data (1955 to 2012) were downloaded as monthly statistical means with a horizontal resolution of 1° (x) x 1° (y) and 57 and 37 vertical (z) levels between 0 to 1500 m and 0 to 500 m for dissolved oxygen and phosphate/nitrate, respectively. The data request was on 11-07-2013 (dissolved oxygen), 17-07-2013 (nitrate) and 18-07-2013 (phosphate), respectively.

2.3.6 Data on chemical sediment components

A data compilation on total organic carbon content and calcium carbonate and silicia in surface sediments were downloaded from the data archive PANGAEA (Seiter et al., 2004a, b, c, and references therein) (see Table 1). Data on biogenic silica of the sediment surface were obtained from PANGAEA, too (see Geibert et al., 2005a, b).

The data described under 2.3.4 to 2.3.6 were used as explanatory variables in the Antarctic krill species distribution model (SDM) (SeaWiFS, WOA13, chemical sediment components) and in the demersal fish SDM (WOA13, chemical sediment components). The SDMs are described in detail in the Supplement and the PANGAEA link to the respective data layer products (incl. file names) is given in the corresponding subsection under "2.4 Ecological data".

2.4 Ecological data

In the following, we describe the sources of raw data sets used in the WSMPA planning process and indicate which data layer product was developed on the basis of which raw data sets per higher taxonomic group. In addition, the methods for processing and analysing the data and for developing each data layer are described in detail in the Supplement.

2.4.1 Zooplankton

Antarctic krill (adults)
The WSMPA data collection on adult Antarctic krill (*Euphausia superba*) originates from (i) historical UK data from “Discovery Expeditions” (1928-1939) and data collected during the SIBEX cruise by British Antarctic Survey, (ii) five South African data sets from the 1990s, (iii) four Soviet data sets from 1998 and 1990, (iv) Polish data (Witek et al., 1985) and (v) German data from location discovery cruises with MV “Polarsirkel” in 1979/80 and 1980/81 (Siegel, 1982), RV “Walther Herwig” cruises (1975/76, 1977/78) and the 2004 Lazarev Sea Krill Survey (LAKRIS) (RV “Polarstern” cruise ANT-XXI/4) (Siegel, 2012). All the data are publicly available via the database KRILLBASE (doi.org/brg8) (Atkinson et al., 2017) (see Table S2 in the Supplement that provides a detailed list of data used from KRILLBASE). The data from KRILLBASE were complemented by abundance data on *E. superba*, which were collected (a) during the Norwegian Antarctic research expedition 1976/77 (MV “Polarsirkel”) (Fevolden, 1979), (b) during two Soviet research cruises in 1977 (RV “Gizhiga”) and 1983 (RV “Volvik Verter”), (c) in the context of the Lazarev Sea Krill Survey (RV “Polarstern” cruises ANT-XXIII/2, ANT-XXII/6, ANT-XXIV/2) (e.g. Siegel, 2012) as well as (d) during RV “Polarstern” cruises ANT-V/1-3, ANT-VII/4, ANT-VIII/4 and ANT-XXIX/3 (Siegel et al., 2013) (Table 2). Furthermore, Japanese, Norwegian and Soviet fisheries data (catch and effort) on *E. superba* for the WSMPA Planning Area (i.e. Statistical Subarea 48.5 and southern part of Subarea 48.6) stem from the CCAMLR database (data request 15 through CCAMLR Secretariat: 03-10-2013) (Table 2).

All these data were used in a species distribution model (SDM) of adult Antarctic krill and ultimately led to a data layer product showing habitat suitability for adult Antarctic krill in the WSMPA Planning Area (see doi 10.1594/PANGAEA.899667; file name: "Adult Antarctic krill, *Euphausia superba* - habitat suitability prediction").

**Antarctic krill (larvae)**

Abundance data on Antarctic krill larvae stem from (a) the Antarctic research expeditions 1976/77 (Fevolden, 1979) and 1979/80 with MV “Polarsirkel” (Siegel, 1982), (b) the First International BIOMASS Experiment survey (FIBEX), (c) (RV “Walther Herwig” cruise 1981) (e.g. Trathan and Everson, 1994) and the Lazarev Sea Krill Survey (LAKRIS) (RV “Polarstern” cruises ANT-XXI/4, ANT-XXIII/2) (Siegel, 2012) as well as (c2) RV “Polarstern” cruise ANT-VII/4 and the combined RV “Polarstern” (ANT-VII/2) and RV “Akademik Fedorov” cruises (Menshenina, 1992) (see Table 2). All data on Antarctic krill larvae were used for an interpolation approach and led to a map of the interpolated abundances of krill larvae in the WSMPA Planning Area (see doi 10.1594/PANGAEA.899667; "Antarctic krill larvae, *Euphausia superba* - interpolated abundance").

**Ice krill**

Abundance data on adult ice krill (*Euphausia crystallorophias*) originate from pelagic trawl surveys during (a) the German Antarctic research cruise 1975/76 with “Walther Herwig”, (b) the “Pre-Site Survey” 1979/80 with MV “Polarsirkel” (Siegel, 1982), (c) the Lazarev Sea Krill Survey (RV “Polarstern” cruises ANT-XXI/4, ANT-XXIII/2, ANT-XXIV/2) (e.g. Siegel, 2012) as well as (d) RV “Polarstern” cruises ANT-V/1-3, ANT-VII/4 and ANT-XXIX/3 (Siegel et al., 2013) (Table 2). The abundance data on *E. crystallorophias* were used for an interpolation approach and led to a map showing the interpolated abundances of ice krill (see doi 10.1594/PANGAEA.899667; "Ice krill, *Euphausia crystallorophias* - interpolated density"). In addition, the abundance data on *E. crystallorophias* were used for "ground truthing"
All data about *E. superba* and *E. crystallorophias*, which were used additionally to KRILLBASE and the CCAMLR database, are stored in the data warehouse of the Thuenen Institute of Sea Fisheries ([https://www.thuenen.de](https://www.thuenen.de); contact: Lara Kim Hühnerlage; kim.huenerlage@thuenen.de) and can be requested on demand.

### 2.4.2 Zoobenthos

**Sponges**

Abundance data and semi-quantitative data on sponges (higher taxonomic groups), which were compiled in the context of the WSMPA planning initiative, originate from zoobenthos data sets. The abundance data are publically available via PANGAEA (see Gerdes, 2014 a-o). The semi-quantitative data set can be requested from us if required (contact: Katharina Teschke, AWI) and is available as presence-absence data set in PANGAEA (see Teschke and Brey, 2019a) (see Table 2). Based on these data, we developed a map of the occurrence of sponges in the WSMPA Planning Area (doi 10.1594/PANGAEA.899645; "Sponges, Porifera - interpolated presence").

**Echinoderms**

The data set on echinoderms consists of presence-absence data on species level for asteroids, abundance data on ophiurid taxa as well as holothurian taxa. The first two data sets are available in PANGAEA (Teschke and Brey, 2019b, c), the latter in the information system biodiversity.aq (Gutt et al., 2014). Publications, which have used these primary data sets, are e.g. Dahms (1996), Gutt (1988) and Gerdes et al. (1992).

These data were used in a clustering approach to ultimately identify the potential habitat for echinoderms in the WSMPA Planning Area by environmental proxies (doi 10.1594/PANGAEA.899645; "Special echinoderm assemblage - pot habitat").

### 2.4.3 Fish

**Antarctic silverfish (larvae and adults)**

The WSMPA data collection on Antarctic silverfish larvae (*Pleuragramma antarctica*) originates from quantitative zooplankton data sets obtained during the RV “Polarstern” cruises ANT-I/2 (Boysen-Ennen and Piatkowski, 1988) and ANT-III/3 (Hubold et al., 1988) and during the Lazarev Sea Krill Survey (LAKRIS) (“Polarstern” cruises: ANT-XXI/4, ANT-XXIII/6, ANT-XXIV/2) (Flores et al., 2014) (Table 2). The first mentioned data are stored in the data warehouse of the Thuenen Institute of Sea Fisheries and can be requested on demand (contact: Lara Kim Hühnerlage). Fish larvae data from ANT-III/3 are available from Hubold et al. (1988) and the LAKRIS data can be requested from Hauke Flores (AWI) if required (see available in PANGAEA (PANGAEA reference will be added during review process).
All abundance data on Antarctic silverfish (adults and larvae) were used for an interpolation approach and led to
a map of the interpolated abundances of *P. antarctica* in the WSMPA Planning Area (doi
10.1594/PANGAEA.899591; "Antarctic silverfish, *Pleuragramma antarctica* - interpolated abundance").

**Demersal fish**

Abundance data on demersal fish and adult *P. antarctica* stem from benthic and pelagic trawl surveys during
seven “Polarstern” cruises between 1996 and 2011 (ANT-XIII/3, ANT-XV/3, ANT-XVII/3, ANT-XIX/5, ANT-
XXI/2, ANT-XXIII/8, ANT-XXVII/3) (Table 2). Publications, which have used these data, are e.g.
Caccavo et al. (2018) and Mintenbeck et al. (2012). The primary data can be requested from us if required
(contact: Rainer Knust, AWI). This data compilation was complemented by data on demersal fish and *P.
antarctica* derived from trawl and dredge surveys published in PANGAEA (Drescher et al., 2012; Ekau et al.,
2012a,b; Hureau et al., 2012; Kock et al., 2012; Wöhrmann et al., 2012).

All data on demersal fish were used in a SDM and led to a data layer product showing the habitat suitability for
demersal fish in the WSMPA Planning Area (see doi 10.1594/PANGAEA.899591; "Demersal fish - habitat
suitability prediction").

**Antarctic toothfish (adults)**

Fishery data (catch per unit effort) on the Antarctic toothfish (*Dissostichus mawsoni*) for the WSMPA Planning
Area (i.e. Statistical Subarea 48.5 and southern part of Subarea 48.6) were taken from the CCAMLR database
and requested through the CCAMLR Secretariat (data request: 03-08-2016) (Table 2).

The data were used to determine the potential habitat of *D. mawsoni* in the WSMPA Planning Area (see doi
10.1594/PANGAEA.899591; "Adult toothfish, *Dissostichus mawsoni* - pot habitat").

**Demersal fish nesting sites**

Information about nesting sites of demersal fish was collected during the RV “Polarstern” cruises PS82 (Knust
and Schröder, 2014) (ANT-XXIX/3) and PS96 (Pippenburg, 2016) (ANT-XXXI/2). The data are available from
Knust and Schröder (2014) (PS82) and Pippenburg (2016) (PS96). The data collected during RV “Polarstern”
cruises were supplemented by data from the literature (Daniels 1978, 1979; Jones & Near 2012). The map with
the locations of the nesting sites of demersal fish is available at PANGAEA (doi 10.1594/PANGAEA.899591;
"Demersal fish - observation of nesting sites") and is also shown in the Supplement (see Fig. S12).

**2.4.4 Flying and non-flying seabirds**

**Breeding and non-breeding Adélie penguins**

Tracking data on breeding and non-breeding Adélie penguins (*Pygoscelis adeliae*) originate from (i) British
Antarctic Survey (BAS) inventory data from Phil Trathan (ID 754) and Mike Dunn and P. Trathan (ID 764 (only
data on breeding Adélies), ID 773, 779), (ii) a data set from BAS (P. Trathan) and Instituto Antártico Argentino
(Mercedes Santos) (ID 753) (Warnick et al., 2019) and (iii) a data set from the US AMLR Program from
Jefferson Hinke and Wayne Trivelpiece (NOAA) (ID 910) (see e.g. Hinke et al. 2015) (see also Table 2). All the
data are stored in the Birdlife International’s Seabird Tracking Database (data request: 20-10-2015). Adélie
penguins breeding locations and estimated abundances of breeding pairs were derived from Lynch and LaRue
(2014).

The tracking data on *P. adeliae* were used to model the probability of breeding and non-breeding *P. adeliae* occurrence during foraging (doi 10.1594/PANGAEA.899520; "Breeding Adélie penguin, *Pygoscelis adeliae* - modelled foraging trips" and "Non-breeding Adélie penguin, *Pygoscelis adeliae* - modelled foraging trips"). The final data layer product for breeding *P. adeliae* also depict breeding locations and estimated abundances of breeding pairs as well as buffer areas around each colony.

Tracking data on non-breeding *P. adeliae* were acquired from BirdLife International`s Seabird Tracking Database, too (data request: 20-10-2015) (Table 2). Downloaded data include (i) BAS inventory data from Phil Trathan (ID 754) and Mike Dunn and P. Trathan (ID 773, 779), (ii) a data set from BAS (P. Trathan) and Instituto Antártico Argentino (Mercedes Santos) (ID 753) and (iii) a data set from the US AMLR Program from Jefferson Hinke and Wayne Trivelpiece (NOAA) (ID 910).

Breeding Emperor penguins

Data on Emperor penguin (*Aptenodytes forsteri*) colony locations and breeding population estimates were derived from Fretwell et al. (2012, 2014) (Table 2). These data were used to develop a probability map of foraging areas for *A. forsteri* (doi 10.1594/PANGAEA.899520; "Breeding emperor penguin, *Aptenodytes forsteri* - modelled foraging areas").

Antarctic petrels

Information on breeding locations and estimated number of breeding pairs of the Antarctic petrel (*Thalassoica antarctica*) were kindly provided by Jan van Franeker (Wageningen University & Research) and are published in van Franeker et al. (1999) (Table 2). The information on breeding pairs and their colony locations is shown in the final data layer product next to modelled foraging habitats of *T. antarctica* (doi 10.1594/PANGAEA.899520; "Antarctic petrel, *Thalassoica antarctica* - modelled foraging areas").

2.4.5 Pinnipeds

Tracking data from pinnipeds were obtained from the MEOP data portal "Marine Mammals Exploring the Oceans Pole to Pole" (data request: 14-11-2016) (see Table 2 for a detailed list of data used). In addition, we have used MEOP data (UK data: ct27, ct70; German data: ct113, wd06, wd07) for which unconditional sharing were not yet accepted at the time of data retrieval and were provided by Lars Boehme (University of St. Andrews) and us (H. Bornemann), respectively. The UK and German data sets are now also freely accessible from the MEOP data portal. Furthermore, the data from the MEOP data portal were complemented by tracking data sets on southern elephant seals (Tosh et al., 2009a, b; James et al., 2012a, b), Weddell seals (McIntyre et al., 2013a, b) and crabeater seals (Nachtsheim et al., 2016a, b) stored in PANGAEA.

All these tracking data were used to model the probability of seal occurrence during foraging (doi 10.1594/PANGAEA.899619; "Seal abundance - modelled prediction values").
Point data from pack-ice seals (unspecified taxa) based on aerial surveys are from Plötz et al. (2011a-e) and were downloaded from PANGAEA (Table 2). These data were sampled during five flight campaigns from 1996 to 2001 within the Antarctic Pack Ice Seals (APIS) programme. In addition, information on crabeater seal densities (predicted or observed) was derived from Bester et al. (1995 and 2002), Flores et al. (2008) and Forcada et al. (2012; Table 2). German and South African APIS data and UK census data were published in e.g. Gurarie et al. (2016) and Forcada et al. (2017), respectively.

All the APIS point data and information on seal densities were used to develop a map showing the distribution patterns of seals in the WSMPA Planning Area (10.1594/PANGAEA.899619; "Seal abundance - modelled and interpolated prediction values").

3 Outlook

This is the first data compilation of data sources for the Antarctic Weddell Sea and adjacent seas, which considers data across the entire ecosystem: i.e., from abiotic data, such as bathymetry and sea ice, to ecological data ranging from zooplankton and zoobenthos to fish, birds and marine mammals. The effort to create such a data compilation of data sources was directly coupled with the initiative to develop a WSMPA. However, our compilation of data sources will facilitate the future research on fauna, ecology and nature conservation in the Weddell Sea. Using our systematic overview of available data for the development of a specific data collection, future projects save the time-consuming multi-parameter data search from the scratch. In addition, our work serves to guide future studies aimed at closing data gaps in the wider Weddell Sea region and/or simply pointing to specific data sets that may be of particular interest to future generations (baseline is a particular issue). For example, however, the data compilation is also suitable for further scientific questions in the wide field of faunistic, ecological and nature conservation studies to investigate the effect of climate change and possible fishing activities in this area. Some of the ecological data sets were collected in the 1980s and earlier, when the Weddell Sea was still almost pristine and hardly affected by any anthropogenic activities, so that these data sets are optimally suited to describe a reference state for assessing the effect of pressures on the Weddell Sea ecosystem. In addition, the ecological data - with a few exceptions - provide information on abundances of the respective taxa and are therefore better suited as an indicator for environmental changes than presence-absence data or presence data only.

Ultimately, the data compilation of data sources serves to protect our data heritage for use by future generations (baseline is a particular issue), to enable work with readily available multi-parameter data sets, and to motivate researchers to add incorporate further data, both from existing "paper sources" and from future measurements, into existing data repositories and archives.

Subsequent work will focus on the development of an efficient and tailor-made management system for the storage of these complex and heterogeneous data and information of WSMPA data compilation and automated data mining, handling and analysis. This system will serve three purposes: (i) to better enable a more holistic and integrative approach towards ecosystem research in the Weddell Sea in general, (ii) to enable the management of the WSMPA to carry out the tasks of the Research and Monitoring Programme as a mandatory part of an MPA under CCAMLR when adopting the MPA, and (iii) to provide key stakeholders and the public with access to data, information and management measures related to the ecosystem of the Weddell Sea region in general and the WSMPA in particular. The CCAMLR MPA Information Repository (CMIR) currently being developed by
the CCAMLR Secretariat will also be available in the future as a suitable storage location for metadata on CCAMLR MPAs in Antarctica.

**Author contribution.** KT collected all data together, described the metadata and led the writing of the paper. HP took over the technical part of the data acquisition (retrieval, storage, processing). VS collected and prepared the data on zooplankton for further analyses within the WSMPA planning. HB and RK were significantly involved in the collection of the data on pinnipeds and fishes, respectively. TB collaborated in the paper writing.

**Competing interests.** The authors declare that they have no conflict of interests.

Acknowledgements. This work was financially supported by the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office for Agriculture and Food (BLE), grant number 2813HS009. In particular, we would like to thank all colleagues from all national and international scientific institutions who have supported us in providing data used to build up the data compilation for the wider Weddell Sea: i.e. from the Alfred Wegener Institute (Hauke Flores, Dieter Gerdes, Julian Gutt, Stefan Hain, Kerstin Jerosch, Rainer Knust, Dieter Piepenburg, Ralf Timmermann), British Antarctic Survey (Phil Trathan), CCAMLR Secretariat (Elanor Miller, Tim Jones, David Ramn), Helmholtz Centre Geesthacht (Verena Haid), Instituto Antártico Argentino (Mercedes Santos), National Oceanic and Atmospheric Administration (Jefferson Hinke), Royal Belgian Institute of Natural Sciences (Anton van de Putte), Stony Brook University (Heather Lynch), Thünen Institute of Sea Fisheries (Karl-Hermann Kock), University of Gothenburg (Tomas Lundälv), University of Padova (Emilio Riginella), University of St. Andrews (Lars Boehme), Wageningen University & Research (Jan van Franeker). The marine mammal data were collected and made freely available by the International MEOP Consortium and the national programs that contribute to it (http://www.meop.net). The seal tracking data ct96 and ct109 are collected by the Integrated Marine Observing System (IMOS). IMOS is a national collaborative research infrastructure, supported by the Australian Government. It is operated by a consortium of institutions as an unincorporated joint venture, with the University of Tasmania as Lead Agent. We thank two anonymous reviewers for careful reading and constructive comments on the manuscript.
Figure 1. CCAMLR Convention Area with its Marine Protected Area (MPA) Planning Domains and the planning area (incl. 200 km wide buffer area near the Antarctic Peninsula) for the development of a MPA in the wider Weddell Sea (red shaded area). Domain 1: Western Peninsula - South Scotia Arc, Domain 2: North Scotia Arc, Domain 3: Weddell Sea, Domain 4: Bouvet Maud, Domain 5: Crozet - del Cano, Domain 6: Kerguelen Plateau, Domain 7: Eastern Antarctica, Domain 8: Ross Sea, Domain 9: Amundsen - Bellingshausen.

Study site in the Antarctic Weddell Sea and adjacent marine regions. Black dashed line indicates the boundaries of Weddell Sea MPA Planning Area including the 200 km wide buffer area near the Antarctic Peninsula. Overview map of the study site in the wider Weddell Sea and its location in the Southern Ocean (top left corner).
Figure 2. Distribution of all data recordings across the wider Weddell Sea region, which were compiled in the context of the WSMPA planning initiative. Figure S1 in the Supplement provides the distribution of data recordings per higher taxonomic group, i.e. zooplankton, zoobenthos, fishes, birds and pinnipeds.
Figure 3. Raster data sets of environmental parameters, which have been used as basic data in a regionalisation analysis of environmental provinces in the context of the WSMPA planning. IBCSO bathymetry (a), AMSR-E sea ice maps (exemplarily for 15 December 2009) (b), FESOM sea bottom temperature and salinity data (exemplarily for December 2009) (c, d).
### Table 1. Data collection of environmental parameters compiled for the development of a marine protected area (MPA) in the wider Weddell Sea (Antarctica). For each raw data set, the name of the data source, the primary reference, such as the data portal on which the data is publicly accessible or the contact to the respective data provider, as well as examples of publications that have used the respective primary data set are listed. In addition, DOI links to the final WSMPA data layer products is provided, which includes the respective raw data set.

<table>
<thead>
<tr>
<th>Content of Data</th>
<th>Name of data source</th>
<th>Reference to primary data set or data provider</th>
<th>Reference to publications, which have used primary data set (exemplarily)</th>
<th>DOI link to ArcMap packages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Depth</strong></td>
<td>International Bathymetric Chart of the Southern Ocean (IBCSO) Version 1.0</td>
<td>Arndt et al. (2013a) [data request: April 2013]</td>
<td>Arndt et al. (2013b)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899595">https://doi.pangaea.de/10.1594/PANGAEA.899595</a>, <a href="https://doi.pangaea.de/10.1594/PANGAEA.899667">https://doi.pangaea.de/10.1594/PANGAEA.899667</a></td>
</tr>
<tr>
<td><strong>Sea ice concentration</strong></td>
<td>Daily AMSR-E Sea Ice Maps</td>
<td><a href="https://seaice.uni-bremen.de/data/">https://seaice.uni-bremen.de/data/</a></td>
<td>Spren et al. (2008)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899595">https://doi.pangaea.de/10.1594/PANGAEA.899595</a>, <a href="https://doi.pangaea.de/10.1594/PANGAEA.899667">https://doi.pangaea.de/10.1594/PANGAEA.899667</a></td>
</tr>
<tr>
<td><strong>Total organic carbon content</strong></td>
<td></td>
<td>Seiter et al. (2014b, 2004b)</td>
<td>Seiter et al. (2014a, 2004a)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899591">https://doi.pangaea.de/10.1594/PANGAEA.899591</a></td>
</tr>
<tr>
<td><strong>Calcium carbonate, silicia</strong></td>
<td></td>
<td>Seiter et al. (2014c, 2004c)</td>
<td>Seiter et al. (2014a, 2004a)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899591">https://doi.pangaea.de/10.1594/PANGAEA.899591</a></td>
</tr>
<tr>
<td><strong>Biogenic silica</strong></td>
<td></td>
<td>Geibert et al. (2005b)</td>
<td>Geibert et al. (2005a)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899591">https://doi.pangaea.de/10.1594/PANGAEA.899591</a></td>
</tr>
</tbody>
</table>
Table 2. Data collection of ecological parameters compiled for the development of a marine protected area (MPA) in the wider Weddell Sea (Antarctica). For each raw data set, the name of the data source, the primary reference, such as the data portal on which the data is publicly accessible or the contact to the respective data provider, as well as the respective cruise reports and/or examples of publications that have used the respective primary data set are listed. In addition, DOI links to the final WSMPA data layer products is provided, which includes the respective raw data set. The raw data sets are structured according to higher taxonomic groups, i.e. zooplankton, zoobenthos, fishes, birds and pinnipeds. Within each higher taxonomic group, the individual raw data sets are sorted by taxa and the oldest data set is listed first each time.

<table>
<thead>
<tr>
<th>Data content</th>
<th>Name of data source</th>
<th>Reference to primary data set or data provider</th>
<th>Cruise reports</th>
<th>Reference to publications, which have used primary data set (exemplarily)</th>
<th>DOI link to ArcMap packages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zooplankton</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult Antarctic krill (abundances)</td>
<td>MV Polarsirkel 1976/77 Database of Thuenen Institute of Sea Fisheries Contact: Lara Kim Hühnerlage <a href="mailto:kim.huenerlage@thuenen.de">kim.huenerlage@thuenen.de</a></td>
<td>Fevolden (1979)</td>
<td></td>
<td></td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899667">https://doi.pangaea.de/10.1594/PANGAEA.899667</a></td>
</tr>
<tr>
<td>Adult Antarctic krill (abundances)</td>
<td>Soviet cruises: RV Gizhiga 1977 and RV Volny Vetter 1983 Database of Thuenen Institute of Sea Fisheries</td>
<td></td>
<td></td>
<td></td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899667">https://doi.pangaea.de/10.1594/PANGAEA.899667</a></td>
</tr>
<tr>
<td>Adult Antarctic krill (abundances)</td>
<td>ANT-XVIII/4 Database of Thuenen Institute of Sea Fisheries</td>
<td>Fahrbach et al. (2003)</td>
<td></td>
<td></td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899667">https://doi.pangaea.de/10.1594/PANGAEA.899667</a></td>
</tr>
<tr>
<td>Adult Antarctic krill &amp; ice krill (abundances)</td>
<td>ANT-XXIX/3 Database of Thuenen Institute of Sea Fisheries</td>
<td>Gutt (2013)</td>
<td>Siegel et al. (2013)</td>
<td></td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899667">https://doi.pangaea.de/10.1594/PANGAEA.899667</a></td>
</tr>
<tr>
<td>Adult Antarctic krill &amp; ice krill (abundances)</td>
<td>ANT-V/3 Database of Thuenen Institute of Sea Fisheries</td>
<td>Schnack-Schiel (1987)</td>
<td></td>
<td></td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899667">https://doi.pangaea.de/10.1594/PANGAEA.899667</a></td>
</tr>
</tbody>
</table>
Adult Antarctic krill & ice krill (abundances)
ANT-VII/4
Database of Thuenen Institute of Sea Fisheries
Arntz et al. (1990)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Adult ice krill (abundances)
RV Walther Herwig 1975/76
Database of Thuenen Institute of Sea Fisheries
Siegel (1982)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Adult ice krill (abundances)
MV Polarsirkel 1979/80
Database of Thuenen Institute of Sea Fisheries
Fevolden (1979)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Larval Antarctic krill (abundances)
MV Polarsirkel 1976/77
Database of Thuenen Institute of Sea Fisheries
Hempel et al. (1983)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Larval Antarctic krill (abundances)
MV Polarsirkel 1979/80
Database of Thuenen Institute of Sea Fisheries
Sieg (1992)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Larval Antarctic krill (abundances)
First International BIOMASS Experiment survey (FIBEX), RV “Walther Herwig” 1981
Database of Thuenen Institute of Sea Fisheries
Fahrbach and Gerdes (1997)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Larval Antarctic krill (abundances)
ANT-VII/4
Database of Thuenen Institute of Sea Fisheries
Arntz et al. (1990)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Larval Antarctic krill (abundances)
ANT-VIII/2 and RV Akademik Fedorov, 1989
Database of Thuenen Institute of Sea Fisheries
Augstein et al. (1991)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Larval Antarctic krill (abundances)
Lazarev Sea Krill Survey (LAKRIS) data (ANT-XXI/4, ANT-XXIII/6)
Database of Thuenen Institute of Sea Fisheries
Smetacek et al. (2005)
https://doi.pangaea.de/10.1594/PANGAEA.899667

Zoobenthos

Sponges (abundances)
ANT-III/2, WH85
Gerdes (2014a–o)
Hempel (1985)
Schnack-Schiel (1987)
Függer (1988)
Arntz et al. (1990)
Bathmann et al. (1992)
Spindler et al. (1993)
Arntz and Gutt (1997)
Fahrbach and Gerdes (1997)
Arntz and Gutt (1999)
Arntz and Brey (2001)
Arntz and Brey (2003)
Arntz and Brey (2005)
Gutt (2008)
Knust et al. (2012)
https://doi.pangaea.de/10.1594/PANGAEA.899645

ANT-V/1
ANT-VI/3
ANT-VII/4
ANT-IX/3
ANT-X/3
ANT-XIII/3
ANT-XX/3
ANT-XV/3
ANT-XVII/3
ANT-XIX/3
ANT-XXI/2
ANT-XXIII
ANT-XXVII

ANT-IX/3
ANT-X/3
ANT-XIII/3
ANT-XX/3
ANT-XV/3
ANT-XVII/3
ANT-XIX/3
ANT-XXI/2
ANT-XXIII
ANT-XXVII

https://doi.pangaea.de/10.1594/PANGAEA.899645
<table>
<thead>
<tr>
<th>Category</th>
<th>Sampling Sites</th>
<th>References</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fishes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fish larvae - <em>Pleuragramma</em></td>
<td>ANT-I/2</td>
<td>Database of Thuenen Institute of Sea Fisheries, Drescher et al. (1983), Boysen-Ennen and Piatkowski (1988)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899591">https://doi.pangaea.de/10.1594/PANGAEA.899591</a></td>
</tr>
<tr>
<td>Antarctica (abundances)</td>
<td>ANT-III/3</td>
<td>Hubold et al. (1988), Hempel (1985)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899591">https://doi.pangaea.de/10.1594/PANGAEA.899591</a></td>
</tr>
<tr>
<td>Fish larvae - <em>Pleuragramma</em></td>
<td>Lazarev Sea Krill Survey (LAKRIS) data: ANT-XXI/4, ANT-XXII/4, ANT-XXIV/2</td>
<td>Contact: Hauke Flores (AWI) <a href="mailto:Hauke.Flores@awi.de">Hauke.Flores@awi.de</a>, Smetsack et al. (2005), Bathmann (2008, 2010), Flores et al. (2014)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899591">https://doi.pangaea.de/10.1594/PANGAEA.899591</a></td>
</tr>
</tbody>
</table>
### Pleuragramma antarctica & demersal fishes

<table>
<thead>
<tr>
<th>Abundance</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANT-XII/3</td>
<td>Drescher et al. (2012)</td>
</tr>
<tr>
<td>ANT-XV/3</td>
<td>Ekau et al. (2012a)</td>
</tr>
<tr>
<td>ANT-V/3</td>
<td>Ekau et al. (2012b)</td>
</tr>
<tr>
<td>ANT-VII/4</td>
<td>Hureau et al. (2012)</td>
</tr>
<tr>
<td>ANT-IX/3</td>
<td>Wöhrmann et al. (2012)</td>
</tr>
<tr>
<td>ANT-XXIII/8</td>
<td>Kock et al. (2012)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Abundance</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANT-I/2</td>
<td>Drescher et al. (2012)</td>
</tr>
<tr>
<td>ANT-III/3</td>
<td>Ekau et al. (2012a)</td>
</tr>
<tr>
<td>ANT-V/3</td>
<td>Ekau et al. (2012b)</td>
</tr>
<tr>
<td>ANT-VII/4</td>
<td>Hureau et al. (2012)</td>
</tr>
<tr>
<td>ANT-IX/3</td>
<td>Wöhrmann et al. (2012)</td>
</tr>
<tr>
<td>ANT-XXIII/8</td>
<td>Kock et al. (2012)</td>
</tr>
</tbody>
</table>

### Antarctic toothfish

**Catch per unit effort**

<table>
<thead>
<tr>
<th>Source</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese, Korean, Norwegian and South African fishing data</td>
<td>CCAMLR database; Contact: CCAMLR Secretariat [data request: 03-08-2016]</td>
</tr>
</tbody>
</table>

### Demersal fish nesting sites

| PS82 (ANT-XXIX/9) | Knust and Schröder (2014) |
| PS96 (ANT-XXXI/2) | Schröder (2016) |

### Birds

#### Adélie penguin colonies

<table>
<thead>
<tr>
<th>Source</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>US AMLR Program (ID 910)</td>
<td>Birdlife International’s Seabird Tracking Database [data request: 20-10-2015]</td>
</tr>
<tr>
<td>BAS / Instituto Antártico Argentino (ID 753)</td>
<td>Birdlife International’s Seabird Tracking Database [data request: 20-10-2015]</td>
</tr>
<tr>
<td>BAS Inventory (754, 773, 779)</td>
<td>Birdlife International’s Seabird Tracking Database [data request: 20-10-2015]</td>
</tr>
<tr>
<td>BAS Inventory (ID 764)</td>
<td>Birdlife International’s Seabird Tracking Database [data request: 20-10-2015]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breeding and non-breeding Adélie penguins (tracking data)</td>
<td>Dunn et al. (2011)</td>
</tr>
</tbody>
</table>

## Contact Information

- Rainer Knust (AWI)
  - Rainer.Knust@awi.de
- Arntz and Gutt (1997)
- Arntz and Gutt (1999)
- Arntz and Gutt (2001)
- Arntz and Gutt (2005)
- Arntz and Brey (2001)
- Arntz and Brey (2003)
- Arntz and Brey (2005)
- Arntz and Brey (2008)

## Additional Resources

- https://doi.pangaea.de/10.1594/PANGAEA.899591
- https://doi.pangaea.de/10.1594/PANGAEA.899552
- https://doi.pangaea.de/10.1594/PANGAEA.899552
- https://doi.pangaea.de/10.1594/PANGAEA.899552
- https://doi.pangaea.de/10.1594/PANGAEA.899552
- https://doi.pangaea.de/10.1594/PANGAEA.899552
- https://doi.pangaea.de/10.1594/PANGAEA.899552
<table>
<thead>
<tr>
<th>Species</th>
<th>Data Source</th>
<th>References</th>
<th>Data Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emperor penguin colonies (populations estimates)</td>
<td>Fretwell et al. (2012)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899520">https://doi.pangaea.de/10.1594/PANGAEA.899520</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fretwell et al. (2014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antarctic petrel Colonies (estimated number of breeding pairs)</td>
<td>Van Franeker et al. (1999)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899520">https://doi.pangaea.de/10.1594/PANGAEA.899520</a></td>
<td></td>
</tr>
<tr>
<td>Pinnipeds</td>
<td>Data from: Australia (ct109, ct96), Brazil (ct56, ct46, ct39, ct22), China (ct105), UK (ct1, ct8, ct27, ct27x, ct40, ct43, ct45, ct49, ct58, ct70), France (ct16, ct62, fl01, fl02, fl11, fl12), Germany (ct21, ct35, ct33b, ct44, ct54, ct68, ct87, ct99, ct102, ct113, wd06, wd07), Norway (ct34), South Africa (ct33, ct50, ct73), USA (ct9, ct14, ct25, ct29, ct37, ct48)</td>
<td>Marine Mammals Exploring the Oceans Pole to Pole (MEOP) data portal [data request: 14-11-2016]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Treasure et al. (2017)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nachtsheim et al. (2019)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Boehme et al. (2016)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899619">https://doi.pangaea.de/10.1594/PANGAEA.899619</a></td>
<td></td>
</tr>
<tr>
<td>Southern elephant seals (tracking data)</td>
<td>Tosh et al. (2009a)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899619">https://doi.pangaea.de/10.1594/PANGAEA.899619</a></td>
<td></td>
</tr>
<tr>
<td></td>
<td>James et al. (2012a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weddell seals (tracking data)</td>
<td>McIntyre et al. (2013a)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899619">https://doi.pangaea.de/10.1594/PANGAEA.899619</a></td>
<td></td>
</tr>
<tr>
<td>Crabbeater seals (tracking data)</td>
<td>Nachtsheim et al. (2016a)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899619">https://doi.pangaea.de/10.1594/PANGAEA.899619</a></td>
<td></td>
</tr>
<tr>
<td>Pack-ice seals (aerial surveys) Antarctic Pack Ice Seals (APIS) programme EMAGE-I to -V</td>
<td>Plötz et al. (2011a-e)</td>
<td>Southwell et al. (2012)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gurarie et al. (2014, 2017a, b)</td>
<td><a href="https://doi.pangaea.de/10.1594/PANGAEA.899619">https://doi.pangaea.de/10.1594/PANGAEA.899619</a></td>
<td></td>
</tr>
</tbody>
</table>
References


CCAMLR-XXXVII/31: Proposal on a conservation measure establishing a marine protected area in the Domain 1 (Western Antarctic Peninsula and South Scotia Arc), Delegations of Argentina and Chile, CAMLR Commission Meeting, Hobart, Australia, 22 October to 2 November 2018, 27 pp., 2018.


Gerdes, D.: Abundance of macrozoobenthos in surface sediments sampled during POLARSTERN cruise ANT-V/1, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, https://doi.org/10.1594/PANGAEA.717708, 2014i.


Pehlke, H., Brey, T., and Teschke, K.: Spatial distribution of a flying seabird (Antarctic petrel) and penguins (Adélie penguin, Emperor penguin) in the wider Weddell Sea (Antarctica) with links to ArcGIS map packages. PANGAEA, https://doi.org/10.1594/PANGAEA.899520, 2019a.


Siegel, V.: Investigations on krill (Euphausia superba) in the southern Weddell Sea, Meeresforsch., 29, 244-252, 1982.


review of data on abundance, trends in abundance, habitat use and diet of ice-breeding seals in the Southern Ocean, CCAMLR Sci., 19, 49-74, 2012.


Teschke, K. and 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 and ANT XVII/3, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, https://doi.pangaea.de/10.1594/PANGAEA.898629, 2019b.

Teschke, K. and 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, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, https://doi.pangaea.de/10.1594/PANGAEA.898773, 2019c.

Teschke, K., Pehlke, H., and Brey, T.: Spatial distribution of demersal and pelagic fishes in the wider Weddell Sea (Antarctica) with links to ArcGIS map packages. PANGAEA, https://doi.org/10.1594/PANGAEA.899591, 2019a.

Teschke, K., Pehlke, H., and Brey, T.: Spatial distribution of zooplankton (Antarctic krill, ice krill) in the wider Weddell Sea (Antarctica) with links to ArcGIS map packages. PANGAEA, https://doi.org/10.1594/PANGAEA.899667, 2019a.

Teschke, K., Pehlke, H., and Brey, T.: Spatial distribution of zoobenthos (sponges, echinoderms) in the wider Weddell Sea (Antarctica) with links to ArcGIS map packages. PANGAEA, https://doi.org/10.1594/PANGAEA.899645, 2019b.


Supplementary Material

S1 Data processing and analysis

S1.1 Environmental data (IBCSO data, AMSR-E sea ice maps, FESOM data)

For our pelagic regionalisation analysis (Fig. S2) we focused on the austral summer (December to March) and used a raster cell size of 6.25 km x 6.25 km (raster size of AMSR-E 89 GHz sea ice concentration maps). For each raster cell (i) the mean of depth (IBCSO data) and depth range (i.e. difference between maximum and minimum depth), (ii) the relative number of days with ice cover ≤ 70 % (AMSR-E sea ice maps) and (iii) the mean of temperature and salinity at the sea surface and the sea bottom (FESOM data) were calculated over the respective time periods of the environmental data sets (detailed description of environmental data sources see main text, Section 2.3, and Tab. 1). The parameters chosen for the pelagic regionalisation analysis are major structuring components of the Weddell Sea ecosystem and are consistent with the variables used by Raymond (2014) in a cluster approach for a circum-antarctic pelagic regionalisation.

For clustering, we used k-means clustering (Han et al., 2011), the most widely used numerical method for partitioning abiotic and/or biotic data in a predefined number of groups (k) (ecological examples from marine realm see e.g., Legendre et al., 2002, Hewitt et al., 2004, Zharikov et al., 2005, Verfaillie et al., 2009). To estimate the optimal number of clusters we used the gap statistic of the R package cluster (Maechler et al., 2014). The first local maximum in the gap statistic was used to define the optimal number of cluster. Due to the large amount of data, the gap statistic could not be applied to the complete data matrix (119,862 samples x 7 variables). Therefore, the matrix was reduced to 4,000 samples x 7 variables by a permutation approach (number of permutations: 150). Finally, the median of the 150 values for optimal number of clusters were used for k-means clustering.

S1.2 Ecological data

S1.2.1 Zooplankton

Antarctic krill (adults)

The habitat suitability model of the Antarctic krill (Fig. S3) was developed with R (R Core Team, 2014) using the biodiversity modelling package biomod2 (Thuiller et al., 2009 and 2014). Biomod is freely available and is probably the best known and most established software in the modelling world of ecologists, geographers and conservationists, combining predictive results from different models (Hao et al., 2018 and references therein).

All models were run with presence-absence data on Antarctic krill (detailed description of data sources see main text, Section 2.4.1, and Tab. 2). The predictor variables used in our final model were defined in a stepwise procedure. First, we fed biomod2 with more than 20 environmental predictors and the model was run. The relative importance of each variable was evaluated by the following permutation procedure: Once the model is calibrated, a standard prediction is generated. Then, one of the predictor variables is randomised and a new prediction is made. The Pearson’s correlation coefficient (r) between that new prediction and the standard prediction is used to measure this variable’s relative importance in the model (= 1 - r; for more details on the permutation procedure see Thuiller et al., 2012). Variables with low importance were then excluded from the subsequent permutation, and the relative importance in the model of each remaining variable was measured again. Based on this permutation procedure (10 permutations in total) we reduced the number of variables to the most important predictors without negatively influencing the model performance. Thus, for our final predictive model we used the following five environmental variables (ranked by decreasing mean importance value calculated by biomod2): (i) dissolved oxygen (WOA13 data), (ii) ice coverage (AMSR-E sea ice maps), (iii) temperature (FESOM data), (iv) bathymetry (IBCSO data)
and (v) chlorophyll-a concentration (SeaWiFS data) (detailed description of environmental data sources see main text, Section 2.3, and Tab. 1). All data used in the models came from near the sea surface in austral summer (January to March).

In our modelling approach, we focused on nine commonly used modelling techniques, which include regression, classification and machine learning methods, as described by Elith and Graham (2009): generalised linear model (GLM), generalised boosting model, generalised additive model, classification tree analysis, artificial neural network, surface range envelope, flexible discriminant analysis, multiple adaptive regression splines, random forest). Three evaluation methods, i.e. relative operating characteristic (ROC), true skill statistic (TSS) and accuracy, were used. Each modelling technique was calibrated with 70% of the data (random sample from the total data set) and the remaining 30% of the data were used to evaluate their performances (Thuiller, 2003). In total 270 calibrated models (9 different models x 10 replicates of pseudo-absences x 3 evaluation runs) were used for the model synthesis where the different models were combined into a single ensemble model (EM). For the development of our EM, all models were scaled applying a binomial GLM as implemented in biomod2 to ensure comparable model results. Out of the 270 individual models we selected those models for our EM with a TSS threshold higher than 0.65 (i.e., good prediction accuracy accord to Thuiller et al., 2010). Furthermore, we ground-truthed our EM against krill catch data from CCAMLR (see main text, Section 2.4.1, and Tab. 2) by calculating the percentage of krill catches in the areas with different predicted habitat suitability (high to unsuitable).

**Antarctic krill (larvae)**

The map of the interpolated abundances of krill larvae in the WSMPA Planning Area (see Fig. S4) was done with the ArcGIS spatial analyst in the ArcGIS desktop software suite (ESRI Inc., 2011) using the inverse-distance weighting (IDW) method, one of the most commonly used deterministic models in spatial interpolation (e.g., Lu and Wong, 2008). The interpolation was performed with log-transformed abundance data (detailed description of data sources see main text, Section 2.4.1, Tab. 2). The output cell size (x, y) was set to 1000 m and the distance coefficient power to 2. The interpolated abundances were finally expressed for a radius of 30 km around each data record.

**Ice krill**

The map of the interpolated abundances of ice krill in the WSMPA Planning Area (Fig. S5) was developed in the same way as the Antarctic krill map (see use of interpolation in previous paragraph). For a detailed description of the data sources see Section 2.4.1 and Table 2 in the main text.

The potential ice krill habitat (Fig. S6) was approximated by water depth from 0 m to 550 m (IBCSO data) and mean sea surface temperature ≤ 0°C (FESOM data) (detailed description of environmental data sources see main text, Section 2.3, Tab. 1). The biological characteristics of ice krill were taken from the Biogeographic Atlas of the Southern Ocean (Cuzin-Roudy et al., 2014).

**S1.2.2 Zoobenthos**

**Sponges**

The map of the occurrence of sponges in the WSMPA Planning Area (Fig. S7) was finally also generated using the IDW method (see use of interpolation in Section S1.2.1 "Antarctic krill (larvae)"). The previous data processing focused on the consolidation of two different data sets (one quantitative, one semi-quantitative; for detailed
description of data sources see main text, Section 2.4.2, and Tab. 2). We transformed the quantitative data into the same four-category system as the semi-quantitative data (i.e. absent, rare, common, very common) by creating a Monte Carlo sample using Sobol low-discrepancy sequences to develop a Weibull distribution (n = 10,000,000). Within the Weibull distribution, the following classes were identified (i) class 0 (absent) = 0, (ii) class 1 (rare) = 0 to mean - standard deviation (SD), (iii) class 2 (common) = mean - SD to mean and (iv) class 3 (very common) = mean to mean + SD. The quantitative data were classified according to these classes and merged with the semi-quantitative data. The interpolated data were finally expressed for a 10 nm radius around each data record according to CCAMLR Conservation Measure 22-09 (2012).

Echinoderms

The potential habitat for echinoderms in the WSMPA Planning Area (Fig. S8) was developed with JMP (S.A.S. Institute Inc.) using Ward’s (1963) minimum variance method, which has been widely used for calculating distances between clusters since its first description (examples from marine realm see e.g., Verfaillie et al., 2009, Weise et al., 2010, Neukermans et al., 2016). A cluster analysis with a species x station matrix was performed for Asterioidea, Ophiuroidea and Holothuroidea respectively (detailed description of data sources see main text, Section 2.4.2, and Tab. 2). All species occurred only in two stations or less were excluded from the clustering. The results of the cluster analyses were then linked to various environmental data sets. Water temperature best reflected the occurrence of a particular echinoderm community, and therefore their habitat was approximated by bottom water temperature ≤-1° (FESOM data).

S1.2.3 Fish

Antarctic silverfish (larvae and adults)

The map of the interpolated abundances of Antarctic silverfish in the WSMPA Planning Area (Fig. S9) was developed using the IDW method (see use of interpolation in Section S1.2.1 "Antarctic krill (larvae)"). For a detailed description of the data sources see Section 2.4.1 and Table 2 in the main text. The interpolated data were finally expressed for a 10 nm radius around each data record according to CCAMLR Conservation Measure 22-09 (2012).

Demersal fish

The habitat suitability model of demersal fish (Fig. S10) was developed with R (R Core Team, 2014) using the biodiversity modelling package biomod2 (Thuiller et al., 2009 and 2014). All models were run with presence-absence data on demersal fish (detailed description of data sources see main text, Section 2.4.3, and Tab. 2). The predictor variables that were used for our final predictive model were defined in the same iterative process as described under "Antarctic krill (adults)". The environmental variables finally used for our modelling approach were (with decreasing variable importance): (i) distance to coast, (ii) bathymetry, (iii) calcium carbonate, (iv) broad benthic positioning index, (v) silica, (vi) dissolved oxygen, (vii) biogenic silica, (viii) total organic carbon, (ix) nitrate, (x) salinity, (xi) temperature, (xii) current velocity, (xiii) slope and (xiv) phosphate (detailed description of environmental data sources see Section 2.3 and Tab. 1 in main text). All data used in the models came from the sea bottom in austral summer (January to March). Distance to coast, i.e. the Euclidean distance to the nearest land from each raster cell centroid (cell size: 8.02 km x 8.02 km), was calculated with the GRASS GIS package v.distance (Soimasuo et al., 1994) in QGIS 2.10 "Pisa". The coastline derived from
IBCSO Version 1.0 DBM (Arndt et al., 2013). Slope and broad scale benthic positioning index (BPI) was also derived from IBCSO and were calculated with the Benthic Terrain Modeler Version 3.0 extension for the ArcGIS desktop software suite (ESRI Inc., 2011). For the calculation of the broad scale BPI, the inner radius was set to 5 km and the outer radius to 125 km according to Jerosch et al. (2016).

In the modelling approach, we focused on the same modelling techniques as described under "Antarctic krill (adults)". ROC and TSS were used as evaluation methods. Each modelling technique was calibrated with 70 % of the data (random sample from the total data set) and the remaining 30% of the data were used to evaluate their performances (Thuiller, 2003). A total of 135 calibrated models (9 different models x 3 replicates of pseudo-absences x 5 evaluation runs) were used for the EM synthesis where all models were scaled applying a binomial GLM as implemented in biomod2 to ensure comparable model results. Out of the 135 individual models we selected those models for our EM with a TSS threshold higher than 0.9 (i.e., high or excellent prediction accuracy accord to Thuiller et al., 2010).

**Antarctic toothfish (adults)**

The probability model of Antarctic toothfish occurrence in the WSMPA Planning Area (Fig. S11) was developed as a function of depth as recommended by the CCAMLR Working Group on Ecosystem Monitoring and Management (WG-EMM). Following analytical steps were performed:

(i) We calculated the standard descriptive parameters of catch per unit effort (CPUE) data on the Antarctic toothfish (CCAMLR fisheries data) per depth interval $i$ ($mCPUE_i$) with a depth interval width of 100 m (depth interval mean depth: $0 \text{ m} \leq D_i \leq 2600 \text{ m}$, $D_{i+1} - D_i = 50 \text{ m}$). Depth intervals with less than five CPUE data points were not considered.

(ii) A Monte Carlo sample was built for each depth interval $i$ ($n = 10,000$) by randomly drawn samples from a log-normal distribution with the same mean and standard deviation as the CPUE data in each depth interval.

(iii) Outliers were defined as data points below $Q1 - 3.0 \times \text{IQR}$ or above $Q3 + 3.0 \times \text{IQR}$ per depth interval $i$ where $Q1$ and $Q3$ are the 25% and 75% quartiles, respectively, and IQR is the interquartile range, i.e. the difference between Q1 and Q3. Thus, only extreme data points, that are "far out" (Tukey 1977), were excluded from the subsequent model fit.

(iv) We fitted a 4 parameter Weibull model to the simulated median $mCPUE_i$, per depth interval $i$.

(v) The median water depth (IBCSO data) was calculated for each raster cell of 6.25 km x 6.25 km (raster size of AMSR-E 89 GHz sea ice concentration maps), was assigned to the respective depth interval and the corresponding $mCPUE_i$ value - calculated by the Weibull model for this depth interval - was mapped (detailed description of IBCSO data see main text, Section 2.3, and Tab. 1). Finally, the potential habitat of the Antarctic toothfish was bounded from 550 to 2 000 m according to CCAMLR Conservations Measures and fishing practice as recommended by WG-EMM (WG-EMM-16 report, para. 3.6).
S1.2.4 Flying and non-flying seabirds

Breeding and non-breeding Adélie penguins

The probability of occurrence of breeding and non-breeding Adélie penguins during foraging (Fig. S13, S14) was developed with R (R Core Team, 2014) using the R package crawl (Johnson, 2015). The continuous-time correlated random walk model developed by Johnson et al. (2008) has become established in Antarctic science in recent years in order to estimate more accurately the locations of tracked seabirds and pinnipeds along their trajectory (see e.g., Warwick-Evans et al., 2018 and 2019; Baylis et al., 2019). Here, we used the random walk model to generate predictions of the location of each tracked Adélie individual on an hourly time scale (detailed description of penguin tracking data see main text, Section 2.4.4, and Tab. 2). Raw ARGOS data were first processed by assigning error values to the different ARGOS location quality codes, i.e. location code 3 (= highest accuracy of ARGOS position estimate) was set off against the lowest error value, the highest error was assigned to location code B (= lowest accuracy of ARGOS position estimate). Subsequently, simulated track-lines between the temporally sequenced ARGOS positions or each tracked individual, were generated by the continuous-time correlated random walk model, binned onto a 6.25 km x 6.25 km spatial grid and pooled per raster cell so that the final data layers (one for breeding, one for non-breeding Adélies) identifies the areas that were used most often by tracked Adélies. Buffer areas (i.e., a 50 km buffer and a 50-100 km ring buffer) around each colony - shown on the final map of breeding Adélie penguins - were adopted in accordance with the recommendations of the 2nd international workshop on the identification of CCAMLR MPAs in Planning Domain 1 (WG-EMM-15/42 and references therein).

Breeding Emperor penguins

The probability model of Emperor penguins occurrence during foraging in breeding season (Fig. S15) was developed as a function of distance from colony and colony size (Fretwell et al., 2012 and 2014) as well as sea ice concentration (AMSR-E sea ice maps; detailed description of data source see main text, Section 2.3, and Tab. 1).

Analysis 1: Probability model of penguin occurrence as a function of distance from colony and of colony size

To calculate the distances from colony for foraging, we used a raster grid with a spatial resolution of 6.25 km x 6.25 km (as for sea ice concentration). We calculated the Euclidean distance for each raster pixel centroid \( j \) to each emperor penguin breeding colony \( i \). Thus, the probability of occurrence \( P_{1,i,j} \) of one penguin from colony \( i \) in centroid \( j \) was calculated by the following approximation:

\[
P_{1,i,j} = \left( \frac{1}{\sqrt{\pi}} \right) e^{-\left( \frac{3d_{l,i,j}}{d_{max}} \right)^2} (1)
\]

where \( d_{max} \) is the maximum foraging distance to breeding colony (here \( d_{max} = 190 \) km; derived from Zimmer et al. (2008) and reference therein by mean maximum foraging distance to the colony of male penguins in winter of 106 km (standard deviation (SD) = 28 km) plus three SD, i.e. 106 km + 3*28km = 190 km), and \( d_{l,i,j} \) is the Euclidean distance (in km) between colony \( i \) and centroid \( j \), which was calculated by:

\[
d_{l,i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} - d. ice_{edge_i} (2)
\]
where \( d_{\text{ice edge}} \) is the distance of colony to the shelf ice edge. Distances \( d_{i,j} \leq 0 \) were set to 1. Subsequently, different boundaries of ice shelf edge were adjusted by a 10 km buffer, which was subtracted from the distances \( d_{i,j} \), too, and a reclassification was performed again (\( d_{i,j} \leq 0 \) were set to 1).

Then, the probability of penguin occurrence \( P_{1i,j} \) from colony \( i \) in centroid \( j \) was normalised between 0 and 1 (i.e. \( 0 \leq P_{1i,j} \leq 1 \)). Finally, all \( P_{1i,j} \) were added for each centroid \( j \) and normalised again to a range between 0 and 1:

\[
P_{1j} = \frac{\sum_{i=1}^{n} P_{1i,j}}{\max(\sum_{i=1}^{n} P_{1i,j})}
\]

where \( n \) is the number of emperor penguin breeding colonies.

To account for breeding colony size (number of animals), each probability of penguin occurrence \( P_{1i,j} \) was weighted with the best population estimate (BE) for this emperor penguin colony according to Fretwell et al. (2012).

\[
P_{1'i,j} = P_{1i,j} \times BE_i
\]

Subsequently, all \( P_{1'i,j} \) were added for each centroid \( j \) and normalised to a range between 0 and 1 (i.e. \( 0 \leq P_{1'i,j} \leq 1 \)):

\[
P_{1'j} = \frac{\sum_{i=1}^{n} P_{1'i,j}}{\max(\sum_{i=1}^{n} P_{1'i,j})}
\]

where \( n \) is the number of emperor penguin breeding colonies.

**Analysis 2: Probability model of penguin occurrence as a function of sea ice concentration**

The probability model of penguin occurrence as a function of sea ice concentration was calculated in the following steps: (1) A sigmoid transfer function was applied (eq. 6) to achieve an even distribution of the mean sea ice concentration data; (2) the ice index data (\( IC_j \)) were normalised to a range between 0 and 1 (eq. 7); and (3) the probability of penguin occurrence was calculated using the transformed data and a hyperbolic \( \tanh \)-function (eq. 8). The mean sea ice concentration was calculated for the breeding period of emperor penguins (June to January) from 2002 to 2011.

\[
I_{C_j} = \frac{1}{1+e^{(-\ln(x \times 10^{-5}) \times \text{gain})}}
\]

with \( x = \) mean sea ice concentration/100 and gain set to 6.23.

Subsequently, the ice index data (\( IC_j \)) were normalised to a range between 0 and 1:

\[
I_{C_j} = \frac{I_{C_j} - \min(I_{C_{j1}, I_{C_{j2}}, ..., I_{C_{jn}}})}{\max(I_{C_{j1}, I_{C_{j2}}, ..., I_{C_{jn}}}) - \min(I_{C_{j1}, I_{C_{j2}}, ..., I_{C_{jn}}})}
\]
For the probability model of penguin occurrence we have assumed that the penguin preference does not relate linearly to sea ice conditions, but with a sigmoid pattern, i.e. areas with medium sea ice concentration are already suitable foraging grounds. This sigmoid pattern was modelled by the following \( \tanh \)-function:

\[
P_{2j} = \frac{\tanh(\pi \cdot (c_{2j} - 1))^2 + 1}{2}
\]

(8)

**Analysis 3: Combining the distance/colony size model with the sea ice concentration model**

An overall probability of penguin occurrence \( P_j \), i.e. a combination of the distance/colony size model and the sea ice concentration model, was calculated by the following equation:

\[
P_j = \frac{\min(P_1, P_2) \cdot \min(P_1, P_2) \cdot \min(P_1, P_2) \cdot \min(P_1, P_2)}{\max(P_1, P_2) \cdot \max(P_1, P_2) \cdot \max(P_1, P_2) \cdot \max(P_1, P_2)}
\]

(9)

**Antarctic petrel**

The potential foraging habitats of the Antarctic petrel (Fig. S16) was developed as a function of (i) sea ice concentration (AMSR-E sea ice maps), (ii) bathymetry (IBCSO data) and (iii) sea water temperature (FESOM data; detailed description of data sources see main text, Section 2.3, and Tab. 1).

As preferred ice regime of the Antarctic petrel we focused on the marginal ice zone, i.e. 15 % - 80 % ice coverage, according to van Franeker (1996) and Ainley et al. (1984, 1994). Data on sea ice concentration were reclassified as first step, i.e. a value of 1 was assigned to each cell with ice cover 15 % - 80 %, whereas cells with ice cover less than 15 % and more than 80 % were set to 0. Then, for each grid cell, the relative number of days (in %) for which a given grid cell had an ice cover between 15 % and 80 % was calculated for the breeding period (January to March) from 2002 to 2011. Subsequently, eight classes regarding the frequency of occurrence of the marginal ice zone were defined and scaled between 0 and 1.

We used abundance data from Ainley and Jacobs (1981) and calculated mean Antarctic petrel densities for three depth classes, i.e. (1) deep ocean: > 2600 m, (2) continental slope and shelf break: 2600 to < 600 m and (3) continental shelf: the remainder of the continental shelf. Then, the mean densities were scaled between 0 and 1.

Finally, bathymetric data (IBCSO) were used to identify the three different depth zones in the Weddell Sea Planning Area.

According to Ainley et al. (1984) Antarctic petrels seem to prefer water temperatures colder than 0.5°C. Thus, sea surface temperature (SST) data (FESOM) were reclassified for each raster cell, i.e. value 3 = SST ≤ 0.5°C in all three months (January to March), 2 = SST ≤ 0.5°C in only two months, 1 = SST ≤ 0.5°C in only one month and 0 = SST > 0.5°C in all three months. Subsequently, the values were scaled between 0 and 1.

Finally, we approximated the potential foraging habitat of Antarctic petrel by stacking the three environmental proxies and corresponding data layers, respectively, and assigning different weighting factors to the proxies. The highest weighting factor was assigned to sea ice concentration (weighting factor: 1) as we assume sea ice as the major structuring component of the Antarctic petrel foraging habitat. Bathymetry and sea water temperature, in contrast, got lower weighting factors of 0.75 and 0.25, respectively.
Subsequently, we combined our model approach with the model results from Descamps et al. (2016) as recommended by the CCAMLR Scientific Committee (SC-CAMLR-XXXV report, paras. 5.14 - 5.28). Descamps et al. (2016) kindly provided us with the shape files showing the modelled kernel utilization summer and winter distribution of Antarctic petrel breeding at Svarthamaren.

We combined the kernel utilization distribution (hereafter kernel UD) model from Descamps et al. (2016) with our model by the following procedure:

(i) We calculated a weighting factor $w_f_i$ for each level of kernel UD (i.e. for 30, 60 and 95 % kernel UD) by the following equation:

$$w_f_i = \frac{\text{max}(k_{UD})}{k_{UDi}}$$

where $\text{max}(k_{UD})$ is 30 derived from the 30 % kernel UD, i.e. core area - high intensity of use, and $k_{UDi}$ is the respective kernel UD.

(ii) We computed the probability of Antarctic petrel occurrence $P_i$ for each grid cell (i) by:

$$P_i = \frac{x_i \text{AWI}_\text{model} \cdot (100 \cdot w_f_i \text{Descamp et al.}_\text{summer}) \cdot (100 \cdot w_f_i \text{Descamp et al.}_\text{winter})}{100 \cdot \text{max}(x_i \text{AWI}_\text{model}, 100 \cdot w_f_i \text{Descamp et al.}_\text{summer}, 100 \cdot w_f_i \text{Descamp et al.}_\text{winter})}$$

where $x_i \text{AWI}_\text{model}$ is our model value (i.e. 5, 20, 35, 50 or 100).

S1.2.5 Pinnipeds

The probability of pinniped occurrence based on tracking data (Fig. S17) was developed with R (R Core Team, 2014) using the R package crawl (Johnson, 2015; see examples of Antarctic studies using crawl in Section S1.2.4 "Breeding and non-breeding Adélie penguins").

Here, we used the random walk model from Johnson et al. (2008) to generate 100 simulated track-lines between the temporally successive ARGOS positions for each tracking data set on pinnipeds (detailed description of pinniped tracking data see main text, Section 2.4.5, and Tab. 2). Only random track-lines were generated where the maximum speed of a pinniped between successive positions was ≤ 2.5 m s⁻¹. The simulated track-lines were binned onto our standard spatial grid (cell size: 6.25 km x 6.25 km) and pooled per raster cell so that the final data layer identifies the areas that were used most often by tracked pinnipeds.

The map on seal densities in the WSMPA Planning Area (Fig. S18) was developed combining modelled and interpolated densities of seals. Predictive density values on crabeater seals were derived from Flores et al. (2008) and Forcada et al. (2012) and were pooled in case of areas where both studies presented model results. Interpolated densities of seals were derived from APIS point data (unspecified taxa) and observed crabeater seal densities (see Bester et al., 1995, 2002; see detailed data description data in main text, Section 2.4.5, and Tab. 2). From APIS point data, seal densities (i.e. individuals/km²) were calculated using the count method for line transect data (e.g., Bester and Odendaal, 2000, Hedley and Buckland, 2004). We used non-standardised data for the density
calculations as the APIS data set is based on video material, and thus at least observer related factors potentially influencing the probability of animal detection are not relevant to consider. The seal densities from Bester et al. (1995) were averaged over the different sampling dates for each transect, and the densities per sampling zones (inner, middle, outer zone; see Bester et al., 2002) were converted from square nautical mile to square kilometer. Finally, all transects were subdivided in sections of circa 5.5 km according to Bester et al. (2002) using QGIS 2.0 "Dufour" with the QChainage plugin and the density values of the respective transect was assigned to each section for the interpolation approach. We applied the IDW method (see also Section S1.2.1 "Antarctic krill (larvae)") with the output cell size (x, y) of 2000 m and the distance coefficient power of 2. The search radius setting, i.e. the number of points, was set to 10.
Figure S1. Distribution of data recordings per higher taxonomic group, i.e. zooplankton (a), zoobenthos (b), fishes (c), birds (d) and pinnipeds (e), across the wider Weddell Sea region, which were compiled in the context of the WSMPA planning initiative.
Figure S2. Pelagic regionalisation of the WSMPA Planning Area.
Figure S3. Habitat suitability predictions of adult Antarctic krill (*Euphausia superba*) in the WSMPA Planning Area.
Figure S4. Interpolated abundances of Antarctic krill larvae (*Euphausia superba*) in the WSMPA Planning Area.
Figure S5. Interpolated abundances of ice krill (*Euphausia crystallorophias*) in the WSMPA Planning Area.
Figure S6. Potential habitat of ice krill (Euphausia crystallorophias) in the WSMPA Planning Area.
Figure S7. Interpolated occurrences of sponges in the WSMPA Planning Area.
Figure S8. Potential habitat of a special echinoderm assemblage in the WSMPA Planning Area.
Figure S9. Interpolated abundances of Antarctic silverfish (*Pleuragramma antarctica*) in the WSMPA Planning Area.
Figure S10. Habitat suitability predictions of demersal fishes in the WSMPA Planning Area.
Figure S11. Probability model for the potential habitat of Antarctic toothfish (*Dissostichus mawsoni*) in the WSMPA Planning Area.
Figure S12. Nesting sites of demersal fish observed in the WSMPA Planning Area.
Figure S13. Modelled probability of the occurrence of breeding Adélie penguins (*Pygoscelis adeliae*) during foraging in the WSMPA Planning Area.
Figure S14. Modelled probability of the occurrence of non-breeding Adélie penguins (*Pygoscelis adeliae*) during foraging in the WSMPA Planning Area.
Figure S15. Modelled probability of the occurrence of Emperor penguins (*Aptenodytes forsteri*) during foraging in breeding season.
Figure S16. Probability model for the potential habitat of Antarctic petrel (*Thalassoica antarctica*) in the WSMPA Planning Area.
Figure S17. Modelled probability of seal occurrence in the WSMPA Planning Area.
Figure S18. Modelled and interpolated seal abundances in the WSMPA Planning Area.
# S3 Tables

Table S2. Detailed list of adult Antarctic krill (*Euphausia superba*) data with survey name, station number per survey and respective source of data, which were used from the database KRILLBASE within the WSMPA planning initiative.

<table>
<thead>
<tr>
<th>Survey name</th>
<th>Station</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>bas1985sib</td>
<td>4a and 5</td>
<td>UK data (SIBEX cruise)</td>
</tr>
<tr>
<td>epa1993saf</td>
<td>211-216</td>
<td>South African data</td>
</tr>
<tr>
<td>epa1995saf</td>
<td>1-14</td>
<td>South African data</td>
</tr>
<tr>
<td>epa1998saf</td>
<td>27</td>
<td>South African data</td>
</tr>
<tr>
<td>epa1995bon</td>
<td>189</td>
<td>South African data</td>
</tr>
<tr>
<td>epa1996xxx</td>
<td>5 and 6</td>
<td>South African data</td>
</tr>
<tr>
<td>epa1989ikt</td>
<td>7, 8, 10, 11</td>
<td>Soviet data</td>
</tr>
<tr>
<td>epa1990mkx</td>
<td>149-173</td>
<td>Soviet data</td>
</tr>
<tr>
<td>epa1989smt</td>
<td>5, 6, 13, 14</td>
<td>Soviet data</td>
</tr>
<tr>
<td>epa1989ikt</td>
<td>7, 8, 10, 11</td>
<td>Soviet data</td>
</tr>
<tr>
<td>his1928dis-top</td>
<td>46</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1931wsc-top</td>
<td>548-552</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1932dis-str</td>
<td>813a, 815a, 816a, 822a</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1932dis-top</td>
<td>814, 818-820</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1933dis-str</td>
<td>1148a, 1150a, 1151a, 1153a</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1933dis-top</td>
<td>1149, 1152</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1935dis-str</td>
<td>1519a</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1935dis-top</td>
<td>1513-1528</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1935wsc-top</td>
<td>892 and 893</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1937dis-str</td>
<td>2004a</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1937dis-ver</td>
<td>2010a, 2012a</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1939dis-str</td>
<td>2596a, 2598a, 2600a, 2606a-2608a, 2010a</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1939dis-top</td>
<td>2543-2563, 2595, 2599, 2601, 2603, 2604, 2609</td>
<td>UK historical data</td>
</tr>
<tr>
<td>his1939dis-ver</td>
<td>2543a, 2545a, 2547a</td>
<td>UK historical data</td>
</tr>
<tr>
<td>pol1984wit</td>
<td>165 and 168</td>
<td>Polish data; Witek et al. (1985)</td>
</tr>
<tr>
<td>vsi1976xxx</td>
<td>161, 162, 176, 183, 185, 186, 189, 190</td>
<td>German data (RV Walther Herwig cruise 1975/76)</td>
</tr>
<tr>
<td>vsi1978xxx</td>
<td>330-337, 339, 340a, 341b, 342c, 343d, 344c, 345</td>
<td>German data (RV Walther Herwig cruise 1977/78)</td>
</tr>
</tbody>
</table>
S4 References


Response to the interactive comments "An integrated data compilation for the development of a marine protected area in the Weddell Sea" by Katharina Teschke et al. (Earth System Science Data)

Anonymous Referee #1 - https://doi.org/105194/essd-2019-86-RC1

Specific comments

1. "At first reading of title and abstract I was under the impression that a huge data compilation is provided with this paper, while in fact a systematic overview of all the sources to the MPA planning process is given. This should be made more clear right from the begin, so the reader knows what to expect."
   a. We agree with the referee's statement that we have to clarify right from the beginning of the paper that a systematic overview of all the data sources - instead of a huge data compilation - used in the MPA planning process is given. We will change the whole text (incl. title) accordingly.

2. "Usability of data: The authors should elaborate if and how the provided data and information might be used, and/or to which future work they might contribute. Can the provided data layer products directly be used by readers in some way? Is the interested reader invited to build up his own data compilation by using the provided sources?"
   From general comments: "How or if the provided data layer products might be used, without access to the underlying data compilation, is not discussed."
   a. We have mentioned in section "2.2 Data availability" (lines 15-16) that our data layer products can be used for geo-statistical analyses within the framework of MPA planning, among other things. Nevertheless, we will add some more details on data usability in the text in the appropriate sections (e.g., "2.2 Data availability", "3 Outlook").

3. "Methods: ..., but they don't give as much detail on the methods they used to process the data and create the respective data layers. A brief description of the provided data layers is missing in the paper, e.g. in 2.4.1 (Zooplankton) it is not mentioned that via the persistent identifiers the maps with interpolated abundances of the two krill species can be accessed, but also a map with habitat suitability. Another example is chapter 2.4.2 (Zoobenthos), here species level data sources for asteroids, ophiuroids, and holothurians are listed, but the provided map layer shows one polygon only (special echinoderm assemblage). It should be made clear in each section what data products the user can access via the links, and how they are created."
   a. We will follow the referee's suggestion by indicating in each subsection under "2.3 Environmental data" and "2.4 Ecological data" which data layer products can be accessed via the PANGAEA links and by adding as supplementary material the method by which the data were processed and the respective data layer was developed.

4. "In my opinion the paragraphs on data sources could be shortened, as every source is listed also in table 1 and 2. Instead I would prefer to read more about the methods (e.g. models used)."
   a. We agree with this statement, that we could shorten the sections "2.3 Environmental data" and "2.4 Ecological data" by avoiding duplication of information in the text about e.g., references to publications and cruise reports, explicitly listed in the tables, too. We will change the text accordingly.
b. We have already commented on this remark earlier in this reply (see #3a.). We will provide the information about the analytical methods (e.g. models) in the revised version.

Technical corrections

5. "Here it is spoken of five persistent identifiers, provided are six (also in 2.2). Suggest to sort the links according to the structure in the paper (from abiotic to seals)."
   a. We change the text accordingly.
   b. We follow the referee's suggestion by sorting the PANGAEA links according to the structure of section "2.3 Environmental data" and "2.4 Ecological data".

6. "p 5 line 11 f > a) b) c) in italics"
   a. We change the writing style to "normal".

7. "p 5 line 32 Possible to provide more detailed contact than institute webpage?"
   a. We will add more information about the data warehouse of the Thuenen Institute of Sea Fisheries where data on krill are stored.

   a. We include Barthel and Gutt (1992) in the "Reference" section because they are missing there. However, we do not include Timmermann (2013) in the "Reference" section, because the correct reference is Haid and Timmermann (2013), referred to in section "2.3.3 FESOM data" and listed in "References". We change "Seiter et al. 2014a, b, c" in the text and Table 1 to "Seiter et al. 2004a, b, c".

Anonymous Referee #2 - https://doi.org/10.5194/essd-2019-86-RC2

Specific comments

1. "Include further description of Methods used to analyze each data set and to develop each map. This could potentially be done in the paper itself as an Annex or in the Supplement section (including the maps), within the metadata file (adding an easier crosslink to the paper), and/or as a footnote/bigger caption in each available map. CCAMLR Working Groups or Workshops papers such as those submitted to EMM/SAM/WS are not generally available for the general public (login is required) so further information included therein should be available elsewhere for the interested reader."
   From general comments: "However, in my view, some information is missing, in particular related to metadata and methods description."
   a. We agree with the referee’s statement that the methods used to analyse each data set and to develop each map should be available for the interested reader. Therefore, we will describe the methods in the "Supplement" section (including the maps).

2. "In the description of the Methods, it could be good to include how the methods in each case were chosen (e.g. agreed by international community, based on specific paper, etc.) so it adds to the openness and transparency of the process."
a. We agree that a sentence about how the methods were chosen in each case increases the openness and transparency of the MPA planning process. We will add this information to the methods in the "Supplement" section accordingly.

3. "It is not clear why only maps for 2.3 Environmental data are included at the end of the paper. I would suggest including maps (and methods) for 2.4 Biological data as well, for an easier and more comprehensive visualization."
   a. We showed all described data for "2.4 Ecological data" (point data) in Figure 1 and mapped in Figure 2 the raster data for "2.3.1 IBCSO data", "2.3.2 AMSR-E sea ice maps" and "2.3.3 FESOM data". We have refrained from presenting the environmental variables, which have only been used as explanatory variables in species distribution models (despite the description of these variables in the text).
   b. Furthermore, we will provide each data layer product as a map in the "Supplement" section in the revised version (see also remark #1a earlier in this reply).

4. "Avoid duplication of information in the text about data sources, references and cruise reports already included in the tables."
   a. We will change text sections "2.3 Environmental data" and "2.4 Ecological data" by deleting e.g., references to publications and cruise reports, explicitly listed in the tables, too.

5. "Most readers would probably be unfamiliar with CCAMLR. I would suggest adding a few general maps, including the CCAMLR Convention Area and the division in MPA Planning Domains (mentioned in the text) for contextualization."
   a. We will follow the referee's suggestion by adding a map including the CCAMLR Convention Area and the MPA Planning Domains.

6. "In the 3. Outlook section, there is some mentioning to the development of a storage management system for this data. I would suggest also mentioning the CCAMLR MPA Information Repository (CMIR) that is under development by the CCAMLR Secretariat, as an additional suitable storage space."
   a. We will add the information in "3 Outlook" that the CCAMLR MPA Information Repository (CMIR), currently being developed by the CCAMLR Secretariat, will also be available in the future as a suitable storage location for metadata.

Technical corrections
i. "Include CRS and projections information in each metadata file (common and thematic layers) for each shapefile and raster."
   a. We will add CRS and projections information in the revised version under "2.2 Data availability".
   b. Furthermore, CRS and projections information is supplied for each shape and raster file in "Source" under "Layer Properties" if you upload the file in GIS-software or open the ArcMAP packages.
   c. In addition, projections information is named for each map in the legend (see folder "map_png").
ii. "Provide clear cross-reference links between metadata description and available maps (names do not always coincide and it is hard to keep track to which description fits which map)."

a. We follow the referee's suggestion by indicating for each metadata description (under "2.3 Environmental data" and "2.4 Ecological data"), which map was developed and under which file name the respective data layer is stored in PANGAEA.

iii. "If possible, allow for the zip data to keep a clear file name referenced to the data they contain for easier identification when downloaded in folders (in particular for the “Data shapefile raster”)."

a. We have had the file names for zip data folders changed accordingly by the great support of the PANGAEA team. Because if a PANGAEA data entry is registered by a DOI, then - strictly speaking - nothing can be changed in the data publication anymore.

iv. "Map legends in Figure 3 are very hard to read – make sure high definition maps are provided in final draft or make maps bigger."

a. We agree that the map legends in Figure 3 are impossible to read. We change the maps accordingly.

v. "In section 2.2 Data availability, paragraph 10, there is the mention to five persistent identifiers. However, six of those are provided. Be aware that the same happens in the Abstract."

a. We change the text accordingly.