Sample elaborations, data analyses, results and conclusions from the sea-going MOSAiC and SAS expeditions in the Central Arctic Ocean

European Maritime and Fisheries Fund (EMFF)

Written by Pauline Snoeijs-Leijonmalm, Serdar Sakinan, Christian Stranne, Hans Nilsson, Jonas Hentati Sundberg, Barbara Niehoff, Nicole Hildebrandt, Hauke Flores, Fokje Schaafsma, Kim Vane, Martina Vortkamp, Filip Volckaert, Sarah Maes, Marie Verheye, Julek Chawarski, Flor Vermassen, Stefan Bertilsson, Moritz Buck, Prune Leroy, Allison Churcher, John Sundh, Marcin Kierczak

May – 2023

This report should be cited as:

Sample elaborations, data analyses, results and conclusions from the sea-going MOSAiC and SAS expeditions in the Central Arctic Ocean

Final Report

Specific Contract
CINEA/EMFF/2020/3.2.1/01/ SC07/SI2.860330

LEGAL NOTICE

This document has been prepared for the European Commission; however, it only reflects the views of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

More information on the European Union is available on the Internet (https://ec.europa.eu).

TABLE OF CONTENTS

LIST OF TABLES

place in the CAO, but MOSAiC was also partly in the inflow area of Atlantic water to the CAO in Fram Strait and SAS-Oden was also partly in the outflow area of Atlantic water from the on the Greenland shelf. The habitats are: ice (snow, melted ice, brine, melt ponds, and water at the ice-seawater interface), water (seawater between 11 and 4500 m of depth), and deep-sea sediment (only sampled during the SAS-Oden expedition). Taxa ending in "_X" represent hits to references for which no consensus species name could be assigned in the database. 197

LIST OF FIGURES

LIST OF ABBREVIATIONS

Term	Description
12S	mitochondrially encoded 12S ribosomal RNA
^{14}C dating	Radiocarbon dating
AAW	Arctic Atlantic Water
ASV	Amplicon sequence variants
ASW	Arctic Surface Water
AWI	Alfred-Wegener-Institut, Bremerhaven, Germany
AWL	Atlantic Water Layer
BOLD	Barcode of Life Data Systems
BP	Before Present
CAO	Central Arctic Ocean
CAOFA	Central Arctic Ocean Fisheries Agreement
CINEA	European Climate, Infrastructure and Environment Executive Agency
CMIP6	The Coupled Model Intercomparison Project (CMIP) Phase 6
COI	mitochondrial cytochrome c oxidase subunit I
COP	Conference of Parties
CT	Conservative temperature
dAAW	dense Arctic Atlantic Water
DG MARE	Directorate-General for Maritime Affairs and Fisheries
DSL	Deep scattering layer
DVM	diel vertical migration
EASME	Executive Agency for Small and Medium-sized Enterprises (predecessor of CINEA)
EC	European Commission
EEZ	Exclusive Economic Zone
EFICA	European Fisheries Inventory in the Central Arctic Ocean
EMFF	European Maritime and Fisheries Fund
EU	European Union
FA	Fatty Acids
FAA	Fatty Acids Analysis
FATM	Fatty Acids Trophic Markers
FiSCAO	Scientific Experts on Fish Stocks in the Central Arctic Ocean (predecessor of PSCG)
FishCam	Deep-sea video camera system targeting fish, squid and zooplankton
FMR	Field Metabolic Rate
FO	Frequency of occurrence (fish stomach analyses)
FWC	Framework Contract
GSI	Gonadosomatic index
HSI	Hepatosomatic index
IB	Icebreaker
IBCAO	International Bathymetric Chart of the Arctic Ocean
ICES	International Council for Exploration of the Sea
JPSRM	Joint Program of Scientific Research and Monitoring of the CAOFA
KUL	Katholieke Universiteit Leuven, Belgium
LIENSS	Institute Littoral Environnement et Sociétés (La Rochelle, France)
LOKI	Light frame On-site Key species Investigation
MARUM	Center for Marine Environmental Sciences, University Bremen (Germany)
MNI	Minimum Number of Individuals (of prey items per fish stomach)
MOSAiC	Multidisciplinary drifting Observatory for the Study of Arctic Climate
NASC	Nautical Area Scattering Coefficient (acoustics)

NBIS	National Bioinformatics Infrastructure Sweden
NBIS-SU	National Bioinformatics Infrastructure Sweden, Stockholm University
NBIS-UMU	National Bioinformatics Infrastructure Sweden, Umeå University
NBIS-UU	National Bioinformatics Infrastructure Sweden, Uppsala University
NOSAMS	National Ocean Sciences Accelerator Mass Spectrometry facility (USA)
PCR	Polymerase Chain Reaction
PI	Principal Investigator
PSCG	Provisional Scientific Coordinating Group of the CAO Fisheries Agreement
PVC	Polyvinyl chloride
RV	Research Vessel
SAS	Synoptic Arctic Survey
SAS-Oden	Synoptic Arctic Survey expedition with icebreaker Oden
SC03	Specific Contract SC03 (MOSAiC expedition 2019-2020)
SC06	Specific Contract SC06 (SAS-Oden expedition 2021)
SC07	Specific Contract 07 (Sample elaborations, data analyses, results and conclusions)
SCG	Scientific Coordinating Group of the CAOFA
SD	Standard deviation
SI	Stable isotopes
SIA	Stable isotopes analysis
SLU	Swedish Agricultural University
SMC	Swedish Metabolomics Centre in Umeå (Sweden)
SND	Swedish National Data Service
SPRS	Swedish Polar Research secretariat
SU	Stockholm University, Sweden
TS	Target strength (acoustics)
UVP6	Underwater Vision Profiler, Version 6
WMR	Wageningen Marine Research, The Netherlands
WP	Work Package
Yr BP	Years Before Present

ABSTRACT

Sample elaborations, data analyses, results and conclusions from the sea-going MOSAiC and SAS expeditions in the Central Arctic Ocean

As a result of global warming, the marine ecosystem around the North Pole, the Central Arctic Ocean (CAO), is in fast transition from a permanently to a seasonally ice-covered ocean. The sea-ice loss is expected to enable summer access to the CAO for non-icebreaking ships, including fishery vessels, in the near future. However, the lack of knowledge on the CAO ecosystem impedes any assessment of the sustainability of potential future fisheries in the CAO. Taking a precautionary approach, the EU and nine countries in October 2018 signed the Agreement to Prevent Unregulated High Seas Fisheries in the CAO (CAOFA). This agreement entered into force in June 2021 and, a.o., requires the establishment of a 16-year-long Joint Programme for Scientific Research and Monitoring (JPSRM) for collecting new ecosystem data in the CAO. To reduce the existing lack of knowledge, the EFICA Consortium participated in two icebreaker expeditions in the Eurasian Basin of the CAO: the MOSAiC expedition 2019-2020 and the SAS-Oden expedition in summer 2021 for collecting field data. This report contains results of data and sample analyses from these expeditions that are relevant for future fish stock modelling and assessment studies. Specifically, the following questions were answered based on the collected field data: (1) Which species of nekton (fish and squid) occur in the CAO? (2) What is the distribution, abundance and biomass of nekton and zooplankton? (3) What is the ecology of the nekton? (4) What is the carrying capacity of the CAO with respect to nekton? The project was highly exploratory, not only in its uniqueness of analysing ecosystem data from this remote area, but also in the development and adaptation of the applied methods. This innovation of methodology forms the basis of the recommendations for the JPSRM provided in this report.

Élaboration d'échantillons, analyses de données, résultats et conclusions des expéditions en mer MOSAiC et SAS dans l'océan Arctique central

L'écosystème marin des régions du Pôle Nord, l'océan central arctique (CAO), historiquement couvert de glace de façon permanente, subit actuellement une transition rapide vers une glaciation saisonnière du fait du réchauffement climatique. Il est attendu que cette diminution de la banquise entraînera dans un futur proche l'accès estival pour les navires, dont les bateaux de pêche, à l'océan central arctique, sans besoin de recourir à des brises glaces1. Cependant, l'absence de connaissances suffisantes sur l'écosystème de l'CAO empêche toute évaluation de la durabilité de futures pêcheries potentielles en son sein. Partant du principe de précaution, neuf états ainsi que l'Union Européenne ont signé en octobre 2018 le Traité pour la prévention de pêcheries non régulées en haute mer dans l'océan central arctique. Cet accord est entré en vigueur en juin 2021 et, entre autres, nécessite la mise en place d'un programme conjoint de recherche scientifique et de surveillance (JPSRM) de 16 ans pour la collecte de nouvelles données écosystémiques dans l'CAO. Pour réduire le manque de connaissances existant, le Consortium EFICA a participé à deux expéditions brise-glace dans le bassin eurasien du CAO: l'expédition MOSAiC 2019-2020 et l'expédition SAS-Oden à l'été 2021 pour la collecte de données de terrain. Ce rapport contient les résultats des analyses de données et d'échantillons de ces expéditions qui sont pertinents pour les futures études de modélisation et d'évaluation des stocks de poissons. Plus précisément, les questions suivantes ont été répondues sur la base des données de terrain collectées: (1) Quelles espèces de necton (poissons et calmars) se trouvent dans l'CAO? (2) Quelle est la distribution, l'abondance et la biomasse du necton et du zooplancton? (3) Quelle est l'écologie du necton? (4) Quelle est la capacité de charge de l'CAO par rapport au necton? Le projet était hautement exploratoire, non seulement dans son caractère unique d'analyse des données écosystémiques de cette région éloignée, mais aussi dans le développement et l'adaptation des méthodes appliquées. Cette innovation méthodologique constitue la base des recommandations pour le JPSRM présentées dans ce rapport.

EXECUTIVE SUMMARY

The SC07 project

The Framework Contract EASME/EMFF/2018/003 (FWC) between the EU Commission and the European Fisheries Inventory in the Central Arctic Ocean (EFICA) Consortium was signed on 2 May 2019 and ended on 28 May 2023. This FWC included three Specific Contracts for operational ecosystem mapping in the Central Arctic Ocean (CAO): SC03, SC06, and SC07. The aim of the SC07 project was to analyse data and samples that were collected in the CAO for ecosystem mapping during the sea-going international MOSAiC expedition (SC03: 2019-2020) with the German research icebreaker "Polarstern" and the SAS-Oden expedition (SC06: 2021) with the Swedish icebreaker "Oden", and to evaluate how the results might be relevant for future fish stock modelling and assessment in the target area of the "CAO Fisheries Agreement" (CAOFA)[1]. Twenty-two scientists from the EFICA Consortium have been involved in the SC07 project, and additionally seven scientists from the Consortium participated in the field work carried out within the SC03 and SC06 projects.

The data analysed within the SC07 project contributes to a growing collection of ecosystem data from the CAO, including those that will be collected within the Joint Programme for Scientific Research and Monitoring (JPSRM) of the CAO Fisheries Agreement (CAOFA) in 2023-2037. Together, these data will allow increasingly detailed assessment models of fish stocks in the CAO in the years to come. The SC07 project contributes with data covering a large part of the Eurasian Basin, the North Pole Area, the Lomonosov Ridge, a corner of the Makarov Basin, and the Fram Strait Atlantic Gateway. Four overarching research questions were addressed in six work packages (WP 2-7) of the SC07 project. The questions overlapped between the WPs, which created many positive synergies within the project. These research questions were:

(1) Which species of nekton (fish and squid) occur in the CAO?
(2) What is the distribution, abundance and biomass of nekton and zooplankton?
(3) What is the ecology of the nekton?
(4) What is the carrying capacity of the CAO with respect to nekton?

The SC07 project was highly exploratory, not only in its uniqueness of analysing ecosystem data from this remote area, but also in the development and adaptation of the applied methods. Innovation of methodology included, e.g., hydroacoustic analyses in the absence of biological samples by scenario-building, a technical method for automatic detection of pelagic fish and squid in video footage, a new method for radiocarbon dating of fish otoliths in deep-sea sediments, and the development of new bioinformatics pipelines for analysing environmental DNA (eDNA) from metagenomes and metatranscriptomes.

Overview of knowledge and knowledge increase 2019-2023

The work performed within FWC EASME/EMFF/2018/003 has increased the knowledge on the CAO ecosystem in the Eurasian Basin between the start of the MOSAiC expedition in September 2019 and the end of the sample and data analyses of the MOSAiC and SAS-Oden expeditions in April 2023 (**Figure 1**). he knowledge on the distributions of sympagic (ice-associated) and pelagic (water-column) fish and their zooplankton prey has increased for the study area and included a first investigation on seasonality during the year-round MOSAiC expedition. However, the knowledge on benthic (seafloor-associated) fish and fish predators and their stomach contents remains very low as this was not the target of the SC07 project (very few species of economic interest). The

---

[1] Agreement to prevent unregulated high seas fisheries in the Central Arctic Ocean. Official Journal of the European Union L 73, 15.3.2019, pp. 3-8 https://eur-lex.europa.eu/EN/legal-content/summary/agreement-to-prevent-unregulated-high-seas-fisheries-in-the-central-arctic-ocean.html
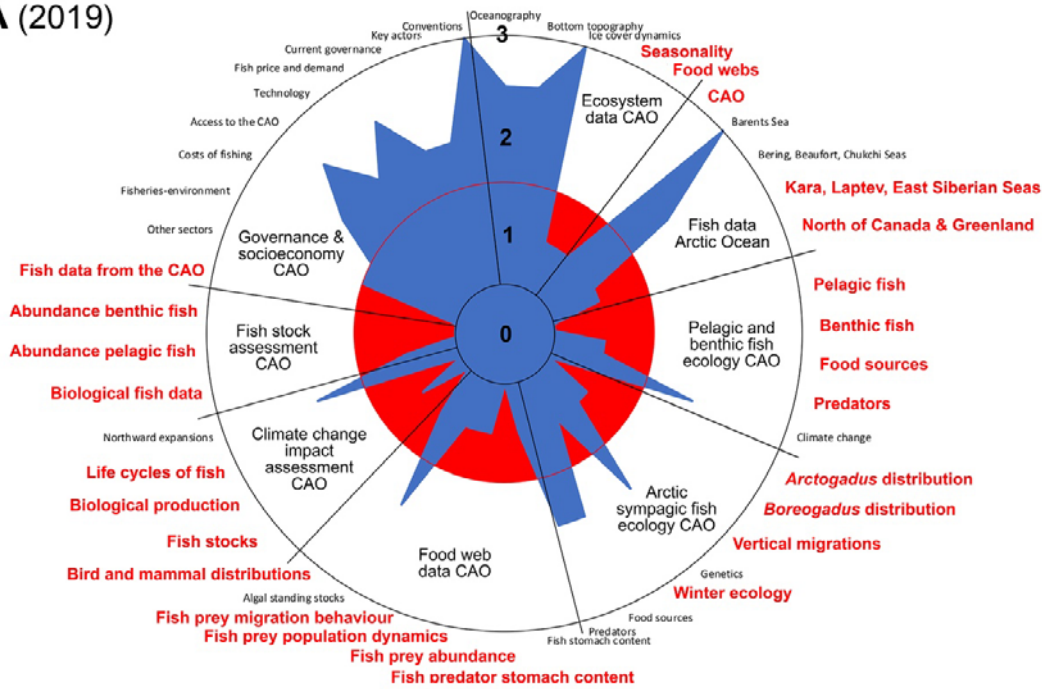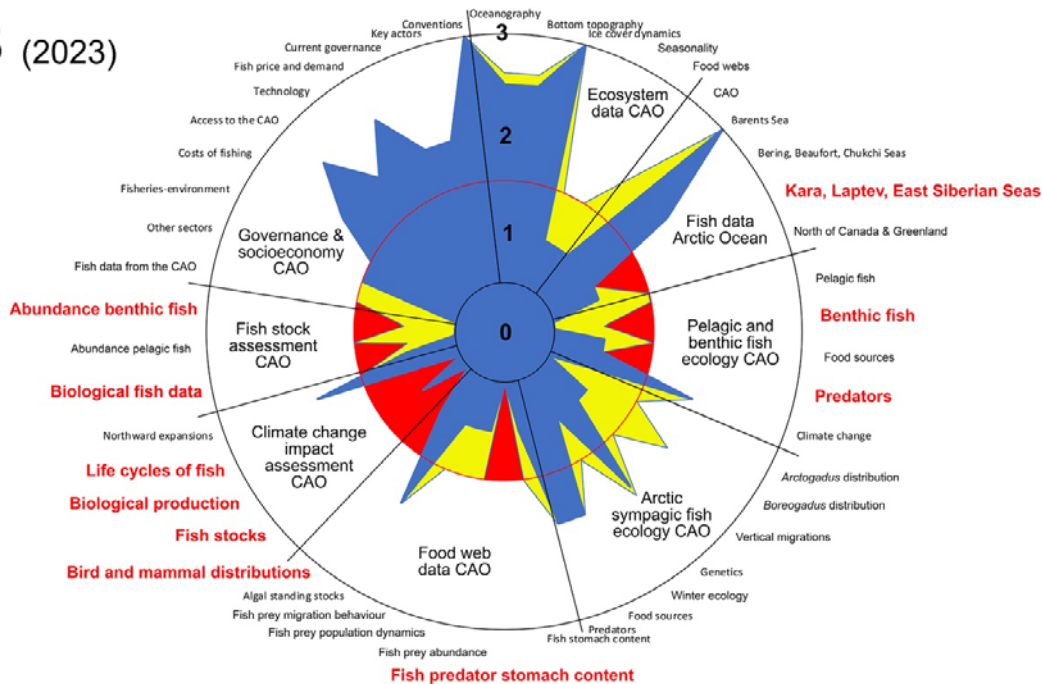
Figure 1: Radar chart summarising knowledge gaps for the Eurasian Basin of the CAO (adapted from Snoeijs-Leijonmalm et al. 2020[2]). (A) Systematically analysed knowledge status September 2019 at the start of the MOSAiC expedition. (B) Estimated knowledge status April 2023 after the sample and data analyses of the MOSAiC and SAS-Oden expeditions. On the axes the estimates of the severity of the knowledge gaps are given: 0 = no knowledge, 1 = serious lack of knowledge, 2 = insufficient knowledge, 3 = sufficient knowledge available for fish stock modelling and assessment. The red circle in combination with the red texts shows where the lack of knowledge is most severe (knowledge status 1.5). The larger the blue area is in the direction of a specific subject, the smaller the relative knowledge gap on this subject. The yellow areas represent the increase of knowledge between September 2019 and April 2023 (but note that the knowledge status was systematically assessed in (A) and only estimated in (B).

[2] Snoeijs-Leijonmalm P, et al. (2020) Review of the research knowledge and gaps on fish populations, fisheries and linked ecosystems in the Central Arctic Ocean (CAO). Publications Office of the European Union [https://data.europa.eu/doi/10.2826/387890]

pelagic knowledge increase is mainly based on new hydroacoustic data, in combination with indirect data (eDNA, sediment otoliths), and a small number of fish samples. Future studies should focus on filling the knowledge gaps by sampling and analysing more fish from the area. Since the focus of the SC07 project was on mapping the ecosystem, climate change impacts were not considered directly. Thus, climate-change effects on biological production, life cycles of fish, fish stocks and bird and mammal distributions are still largely unknown, except for northward expansions of some species that have been observed in the Arctic shelf seas. With respect to the physical environment, the MOSAiC and SAS-Oden expeditions, as well as two Norwegian expeditions (2021, 2022), in the Eurasian Basin have also marginally increased the knowledge on oceanography and bottom topography, which was already relatively high compared with biological knowledge. Ice cover dynamics (open water vs. ice cover), which is followed on a daily basis from satellites, was already high as well.

Table 1: Increase in the knowledge of nekton species of commercial interest in the Eurasian Basin of the Central Arctic Ocean (CAO) provided by the SC07 project and one Norwegian expedition[3] (the latter indicated by *). DSL = deep scattering layer (200-600 m of depth). AGAT = Atlantic Gateway.

Species	Evidence before 2019	Evidence 2019-2022	Distribution
Gonatus fabricii - armhook squid	4 individuals dead in ice holes (one gravid female) at 3 stations, 1 juvenile in plankton	>100 individuals on video, 1 small individual sampled, eDNA	DSL, Siberian side
Boreogadus saida - polar cod	Sympagic juveniles widespread	Sediment otoliths adults Possibly on video, eDNA	Sympagic and DSL
Arctogadus glacialis - ice cod	Sympagic, 3 stations near shelf	Sediment otoliths adults eDNA, 1 sample CAO**	Sympagic and DSL
Gadus morhua - Atlantic cod	-	Video, eDNA 3 samples CAO 11 samples AGAT	DSL
Benthosema glaciale - glacier lantern fish	-	Video, eDNA 3 individuals sampled 2021*	DSL
Gadus chalcogrammus - walleye pollock	-	eDNA	DSL
Melanogrammus aeglefinus - haddock	-	eDNA 38 samples AGAT	DSL
Sebastes mentella - beaked redfish)	-	eDNA 4 samples AGAT	DSL
Reinhardtius hippoglossoides - Greenland halibut	1 juvenile near shelf	eDNA 1 juvenile sampled 2021*	Bentho-pelagic, Greenland side

Nekton species in the CAO

The here reported new results show that the pelagic food web of the CAO ecosystem is more complex than previously thought. The SC07 project increased the number of fish species recorded in the CAO residing in the mesopelagic deep scattering layer (DSL) with three large Atlantic predatory fish: Atlantic cod *Gadus morhua* over the deep Amundsen Basin, and haddock *Melanogrammus aeglefinus* and beaked redfish *Sebastes mentella* at the southern end of the CAO in the inflow of Atlantic water to the Arctic basin (**Table 1**). The DSL resides in the Atlantic Water Layer (AWL) of the CAO with slightly higher water temperature (up to ca. 2 °C) than the water layers above and below which have temperature below 0 °C. The SC07 project detected eDNA of these three large predatory species, as well as of Greenland halibut *Reinhardtius hippoglossoides*, in the CAO as far as the Lomonosov Ridge. This might imply that not only Atlantic cod, but also haddock,

---

[3] Ingvaldsen RB, et al. (2023) Under-ice observations by trawls and multi-frequency acoustics in the Central Arctic Ocean reveals abundance and composition of pelagic fauna. Scientific Reports 13:1000 [https://doi.org/10.1038/s41598-023-27957-x]

beaked redfish and Greenland halibut could occur further north than hitherto known. However, fish samples are needed to confirm this because eDNA can travel far with water currents, especially in cold regions. Fish catches in the CAO included also polar cod *Boreogadus saida* from the sea-ice habitat and another gadoid fish initially identified as ice cod *Arctogadus glacialis* from the DSL. The latter specimen is at present being investigated by fish taxonomists, because genotyping identified it as *Gadus chalcogrammus* (**Table 1**). A deep-sea camera system revealed two more nekton species in the DSL: glacier lanternfish *Benthosema glaciale* and armhook squid *Gonatus fabricii*. The squid was known to occur and even reproduce in the CAO, but not the lanternfish. This observation was later confirmed by a Norwegian expedition in 2021 that caught three individuals of *Benthosema glaciale* (**Table 1**).

The dominance of polar cod *Boreogadus saida* and ice cod *Arctogadus glacialis* otoliths in the sediment samples provides a first indication that, both recently and in the past, no other abundant pelagic fish species would have occurred in the research area of the SAS-Oden expedition. The sediment otoliths provide evidence that *Boreogadus saida*, *Arctogadus glacialis*, and *Paraliparis* spp. have occurred in the CAO during the whole Holocene. The ^{14}C dating results further suggest that these two species may even have been present in the Arctic Ocean during a period of the last ice age approximately 30 000 – 45 000 years before present. This time period roughly compares to published ^{14}C dating results from (wrongly identified) Arctic sediment otoliths[4]. WP6 has used the latest knowledge on accurate dating of biogenic carbonate structures[5] by removing metal deposits, organic material/sediment and autigenous carbonate, and the ^{14}C dating results reported here should be reliable.

Although otoliths from juvenile *Boreogadus saida* (1-2 years old) dominated in the sediment samples, many otoliths of *Boreogadus saida* (up to 5 years) and *Arctogadus glacialis* (up to 12 years) were from adult fish as well. These results suggest that the overall distribution of these species in the CAO may even today contain a higher proportion of older individuals than previously assumed, which would support the hypothesis that juveniles use the CAO sea ice cover as habitat while adults contribute to the mesopelagic deep scattering layer (DSL) that was observed with hydroacoustics in the SC07 project.

Distribution of nekton and zooplankton in the CAO

The results from temperature reconstructions and population genetic analyses carried out within the SC07 project suggested that Atlantic fish species enter the CAO from adjacent European shelf seas. Atlantic cod and haddock originated from stocks in the northern Norwegian Sea and the Barents Sea, including the Svalbard archipelago. Atlantic cod most likely belong to the NEAC stock managed in ICES Areas 1, 2a and 2b whereas haddock belongs to the Northeast Arctic stock managed in ICES Areas 1 and 2b. The Arctic-endemic polar cod *Boreogadus saida* sampled in the CAO originated from spawning populations upstream of the Transpolar Drift. However, a small but significant proportion of the CAO population is possibly advected with the Beaufort Gyre from as far west as the Beaufort Sea.

The hydroacoustic survey performed during the MOSAiC expedition showed that potential fish abundance of both endemic and Atlantic fish was extremely low in the CAO, but that it was unexpectedly high in the inflow of Atlantic water to the Arctic basin in Fram Strait and near the Yermak Plateau in May 2020. However, high abundances at the Yermak Plateau were not observed in September 2021 during the SAS-Oden expedition. When entering the Nansen Basin there was a sharp front from high to low fish density. The overall lowest abundance in this study was recorded at the Gakkel Ridge with minimum

[4] Hillaire-Marcel C, et al. (2022) Challenging radiocarbon chronostratigraphies in central Arctic Ocean sediment. Geophysical Research Letters 49:1-8 [https://doi.org/10.1029/2022GL100446]
[5] Wollenburg JE, et al. (2023) Omnipresent authigenic calcite distorts Arctic radiocarbon chronology. Nature Communications Earth & Environment 4:136 [https://doi.org/10.1038/s43247-023-00802-9]

and maximum numbers of fish observed per day of 1 600 and 4 600 individuals, respectively, while the corresponding values in Fram Strait were 120 000 and 4 430 000 individuals per day. Surprisingly, minimum and maximum numbers per day were slightly higher again in the North Pole / Lomonosov Ridge area north of 88 °N (minimum 13 000 and maximum 138 000 individuals per day) than in the Nansen Basin. Inside the CAO there seemed to be a fish density optimum (DSL) between 1 and 2 °C in the Atlantic Water Layer (AWL). Below 1 °C no larger acoustic targets indicative of Atlantic predatory species were observed.

Due to the lack of biological samples, two scenarios were used to estimate fish biomass in the DSL inside the CAO (excluding Fram Strait). Biomass values for Scenario 1, assuming the occurrence of only the Arctic endemic gadoids *Boreogadus saida* and *Arctogadus glacialis*, varied between 20 kg km^{-2} (Gakkel Ridge) and 22 000 kg km^{-2} (Yermak Plateau), and Scenario 2, assuming higher diversity in an assemblage of the two Arctic endemic gadoids, Atlantic predatory fish, myctophids, and armhook squid, varied between 10 kg km^{-2} (Gakkel Ridge) and 181 000 kg km^{-2} (Yermak Plateau).

Armhook squid was not caught during the two expeditions, except for a single juvenile at the Lomonosov Ridge during MOSAiC in a zooplankton sample, but it was frequently observed on the MOSAiC deep-sea video camera system ("FishCam"). Squid abundance was estimated under completely different conditions than the fish abundance estimations by hydroacoustics. The FishCam was hanging under the ice at a fixed depth, it had a limited field of view (max. ca. 10 m from the camera lens), and squids were obviously attracted by the light. Under these circumstances, more than five squids per day were counted on the Siberian side of the Amundsen Basin in October 2019, which successively decreased to less than one squid per day while the MOSAiC expedition drifted along its route. This distributional pattern was exactly mirrored by the armhook squid's eDNA.

Fish abundance and zooplankton abundance co-varied. For example, the area near the Lomonosov Ridge, where estimated fish biomass was an order of magnitude higher than during the rest of the SAS-Oden expedition, was also characterised by exceptionally high zooplankton biomass. Generally, zooplankton abundance and biomass were higher in the Eurasian Basin, on the Yermak Plateau and on the Lomonosov Ridge than in the Makarov Basin and the Morris Jesup Rise. This general pattern was also reflected in the distribution of hydroacoustic backscatter and fish biomass, except for the Yermak Plateau. Throughout the whole study area of the SAS-Oden expedition, the maximum abundance of macrozooplankton amphipods attracted by the red light red light of a FishCam on the CTD was found at 600-700 m of depth, i.e., at the lower end of the DSL. It could be hypothesised that their deep occurrence is a strategy to avoid predation by fish in the DSL.

Ecology of the nekton

The stomach contents of Atlantic cod, haddock and beaked redfish carried out within the SC07 project indicated that all fish were actively feeding, both in the Atlantic inflow region and in the CAO. Their diet consisted predominantly of macrozooplankton amphipods, especially *Themisto abyssorum* and *Themisto libellula*. On the shallower Yermak Plateau, also benthic prey was taken. The overall contribution of fish and squid in the stomach contents appeared low in terms of numbers, but may be significant in terms of biomass. Both in the Yermak Plateau region and in the CAO, the analysis of condition index, gonadosomatic index and hepatosomatic index indicated that the majority of fish were in good health and had a nutritional value comparable to commercially harvested fish of the same species in the North Atlantic.

Results from carbon and nitrogen stable isotope analysis of fish muscle tissue confirmed that sympagic (juvenile) polar cod were dependent on ice-associated zooplankton such as *Apherusa glacialis* and *Calanus glacialis*. This observation was confirmed by the fatty acid pattern, in which diatom-associated fatty acids, indicative of ice algae, dominated

over flagellate-dominated fatty acids, indicative of phytoplankton. In Atlantic cod, haddock, beaked redfish and ice cod, carbon and nitrogen stable isotope values clustered apart from polar cod and their zooplankton prey in the Arctic Ocean. This could indicate that the time spent in the Arctic Ocean may not have been sufficient to transfer the isotopic signal of the Arctic pelagic food web to their muscle tissue, due to long turnover times. The relative marker fatty acid contributions showed that polar cod and black seasnail were predominantly associated with fatty acids indicative of diatoms. Haddock, Atlantic cod and beaked redfish caught in the Atlantic inflow region were associated with a stronger signal from pelagic flagellate fatty acids. Atlantic cod caught in the CAO showed a higher contribution of diatom-associated markers, indicating a partial contribution of the cold-adapted food web of the CAO. Altogether, the results from this project suggest that pelagic prey in the Arctic Ocean seems to enable survival of Atlantic fish. The potential zooplankton prey species inhabiting the CAO were shown to be of similarly good quality in terms of carbon and nitrogen content and fatty acid composition as in other regions, such as the Barents Sea and the North Atlantic where fish is commercially caught. Thus, it can be concluded that not the food quality but the quantity limits growth and reproduction in fish and other nektonic species such as squid.

Carrying capacity

A comparison of the spatial distribution of fish biomass and zooplankton biomass carried out within the SC07 project suggests that fish distribution generally mirrors the distribution of their zooplankton prey in the CAO. The MOSAiC expedition drifted in both high and low productivity areas, whereas the SAS-Oden expedition was in a low-productivity area only. Since zooplankton species abundance data from the MOSAiC expedition are not yet available (work performed within other projects than the FWC), it is difficult to estimate the amount of fish biomass based on zooplankton biomass. However, applying standard assumptions of trophic transfer theory to the SAS-Oden zooplankton biomass data suggested that the zooplankton community could at maximum support 4 000 kg fish km^{-2}. Fish biomass derived from hydroacoustic measurements during the SAS-Oden expedition was well below this maximum range: the estimated mean pelagic fish biomass density in the water column between the surface and 600 m was between 20 and 230 kg km^{-2}, assuming a community dominated by polar cod and myctophids. This value does not include the biomass of sympagic (juvenile) polar cod dwelling in the under-ice habitat because under-ice prey biomass could not be quantified with the methods applied during the SAS-Oden expedition. A more accurate estimate of the carrying capacity will be possible based on future comprehensive spatial models using zooplankton distribution data from both the MOSAiC and the SAS-Oden expeditions.

Relevance of the collected data for fish stock modelling in the CAO

Habitat modelling: A generalised additive modelling (GAM) approach was used to explore the predictive power of environmental and prey variables to demonstrate the potential habitat preference of the fish communities of both Arctic and Atlantic origin. In this GAM model, the four variables, temperature, latitude, depth and zooplankton abundance had high predictable power for the backscatter from fish, explaining 57.7% of the deviance. These predictions can, after refinement with the help of the MOSAiC zooplankton data, be used to identify areas where future monitoring activities could be recommended to take place. Indications were found that the North Pole area, the Lomonosov Ridge and the eastern Amundsen Basin could be such areas. This can to a large extent be explained by advection patterns as described in an already published paper[6] (which is part of this report). In the future, the data collected in the SC07 project – and upcoming data collected by others – could be used to predict fish abundances in the CAO at a basin-wide

---

[6] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

scale using the relatively large amount of oceanographic and seafloor data available for the CAO. While the seafloor provides refuges of different types for different species, oceanographic data provide circulation patterns and pelagic habitats for fish. Measures of ecosystem productivity such as chlorophyll data, UVP particle distribution data and zooplankton biomass data can be included in models predicting fish productivity.

Food-web modelling: The carbon and nitrogen data obtained within the SC07 project can go straight into food-web models. Because the Arctic ecosystem is changing rapidly, it is important to monitor the long-term variability of the carbon flux, food web structure and nutritional value of zooplankton prey in the CAO in order to assess the future sustainability of potential fisheries. Collection of basic data on stable isotope composition and C:N ratios is a valuable and cost-effective tool to monitor potential changes of the food web and the nutritional value of fish prey during the coming decades.

Population size modelling: The sample size of the Atlantic fish collected by the EFICA Consortium is far too small to estimate population parameters based on genetics. However, the SC07 population genetic data sets can complement other data sets to predict population size in the future. After careful evaluation, it may be possible to estimate effective population size of polar cod drawing on extensive external data sets used to investigate population structure.

Long-term changes in fish species composition with environmental change: In this report it is shown that sediment otoliths can provide important information about the potential long-term species composition, age and size structure of fish in areas where the sampling of living fish is impossible. This is achieved by relating fish age at the time of death derived from the otoliths to radiocarbon dating of the otolith and extrapolating fish size from known age-size relationships for each species. Such knowledge can help to improve fish distribution models and niche modelling. Species identification of sediment otoliths can also aid in understanding the spatio-temporal changes in fish species composition in an area. In the present study it was established that the sediment otolith assemblages in the CAO are dominated by the endemic gadoids *Boreogadus saida* and *Arctogadus glacialis*. Analyses of sediment otoliths can give an indication if a species has occurred in the CAO already for a longer time period, or if it recently expanded its range into the CAO. However, the SC07 data set is yet too small to exclude that Atlantic cod has been living in the CAO for thousands of years or not. No otoliths of Atlantic cod were found in the SC07 project (SAS-Oden only), and future otolith studies on the Siberian side of the Amundsen Basin (where this species was caught during MOSAiC expedition) and the Siberian shelf slope are strongly recommended. Such studies can use new otolith samples and/or already collected geological samples.

Relevance of the collected data for fish stock assessment in the CAO

Stock assessment is a scientific method for successful fisheries management. The aim is to achieve a sustainable level of harvest that maintains the health of the fish population while maximizing the economic gain. The main objective of the stock assessment process is to predict future abundances of the stock from the perspective of different fisheries management strategies by testing different harvest scenarios. Primary needs for an assessment is data from a broad range of parameters, such as age structure of the stock, spawning patterns, fecundity, sex ratio, mortality, growth rate, migratory patterns, food and habitat preferences and the biomass. In order to make reliable projections, these parameters need to be monitored at high spatial resolution in large areas over many years. The assessment models become reliable when there is a long enough time series.

For obvious reasons, most notably the lack of biological samples representative of the fish stocks in the CAO and long-term measuring series, the SC07 project results cannot be directly used for stock assessment at this stage. Instead, the contribution of the project consists of providing – for the first time – a realistic range of species, size and

biomass distribution over a large geographical area of the CAO, thereby constituting a strong reference point for future studies.

Recommendations for the Joint Program for Scientific Research and Monitoring (JPSRM) of the Agreement

The principal recommendation for the JPSRM based on the SC03, SC06, and SC07 projects is that standardization of sampling and analytical methods, and strong interdisciplinary and international collaboration, will be of utmost importance for generating usable data. The EFICA Consortium scientists have learned to standardise sampling methods as much as possible by combining two very different expeditions: MOSAiC being a drift expedition (a "one-year long ice station") and the SAS-Oden expedition trying to collect as many ecosystem parameters a possible while steaming as fast as possible between sampling stations. The interdisciplinary and international collaborations within the SC03, SC06, and SC07 projects have greatly contributed to the success of the projects. It is foreseen that joint ecosystem-based expeditions in the high seas of the CAO within the framework of the JPSRM will be both challenging and enriching.

Summary of recommendations for the JPSRM

1. *Define standard JPSRM data collection settings for ship-mounted echosounders.* To get the best possible data the echosounder should be dedicated to fish and zooplankton targets. Transducers should be at least 38 kHz (standard for fish), 18 kHz (deeper signals possible) and 200 kHz (standard for mesozooplankton).

2. *All noise disturbances of the acoustic data should be avoided as much as possible.* There are many sources of noise that come from the ship and this should be tested thoroughly and alleviated before leaving for a long expedition to the CAO. Noise from ice-breaking is unavoidable, but it is possible to make regular stops, turn off the ship's engines and collect good acoustic data for 10 minutes (e.g., every hour). Drift expeditions are best for acoustic surveys (no ice-breaking, no engines on). Acoustic measurements away from the ship, e.g., WBAT by helicopter can also be a solution.

3. *Develop a standard "JPSRM CTD package".* This costs no extra ship time and always provides salinity, temperature, depth, oxygen, fluorescence together with the JPSRM measurements. The list below contains medium-expensive equipment but very low costs compared to what it costs to go on an expedition to the CAO. The equipment should be well-calibrated and ready well in advance of the expedition. Back-up equipment is mandatory – motherboards are sensitive and can usually not be repaired on board. The standard CTD package is recommended to include:

   * WBAT 200 or 333 kHz (zooplankton backscatter)
   * WBAT 38 kHz (fish backscatter, but only if no ship-mounted echosounder is available)
   * Setup of a very precise accelerometer and compass, or a CTD stabiliser preventing tilt and rotation
   * Deep-sea camera connected to the ship with a fibreoptic cable
   * Light meter
   * Chlorophyll fluorometer (ecosystem productivity)
   * UVP (ecosystem productivity & optical zooplankton composition)
   * Large rosette for water sampling for eDNA (e.g., 24 Niskin bottles of 12 L each)

4. *Develop a standard "JPSRM set-up" for on-board measurements of zooplankton acoustic properties.* These measurements needs to be performed when the zooplankton organisms are alive. Accurate knowledge on macrozooplankton abundance is important for predicting the potential preferred fish habitat.

5. *Deploy "JPSRM ice-tethered autonomous acoustic buoys" that can transmit data to land.* This is a relatively cost-effective way to get a lot of acoustic data from remote areas[7].

6. *Deploy a deep-sea camera system in the DSL.* The system should be rated to at least 500 m with electricity and fibre-optic cable, preferably with cameras at different depths in the DSL. This is only worthwhile on drift expeditions / ice stations of at least three weeks due to very low abundances of organisms. Baited camera systems could be tested.

7. *Design a standard "JPSRM longline".* Longline fishing has proven to be a reliable tool. It should be used only when targets with strong backscatter are observed on the ship's echosounder. This method targets large predatory fish such as Atlantic cod, which could otherwise not be obtained. Perhaps lines can be designed to target smaller fish as well, but during the MOSAiC and SAS-Oden expeditions acoustic observations showed that especially the smaller fish fled as soon as something was lowered in the water column.

8. *Deploy a standard mid-water trawl.* This can be used ad-hoc when open water is available[8]. Patches of open water and even broader leads can occur between ice-floes due to wind forcing in the CAO, but their occurrence is unpredictable and often of short duration. Examples of equipment that can be used are Tucker trawl, RMT (Rectangular Mid-water Trawl) and IKMT (Isaacs-Kidd Midwater Trawl Net). This method targets mesopelagic macrozooplankton and smaller fish (polar cod, myctophids).

9. *Deploy a Surface and Under-Ice Trawl (SUIT).* This trawl was designed for under-ice fish (juvenile polar cod) and ice invertebrates[9].

10. *Design a standard "JPSRM macrozooplankton light-trap line".* In the SC07 project it was shown that macrozooplankton species, especially amphipods, are major food items for the mesopelagic fish in the CAO, but these larger animals are rarely caught in multinet samples (targeting mesozooplankton). To estimate access to food for mesopelagic fish, it would be highly recommended to develop a standard light-trap line to be deployed at standard depths in different areas of the CAO for a specific period of time. The macrozooplankton is attracted fast by light (scale of minutes) and deployment time could be one to several hours, but should be tested, The line should cover at least 0-800 m of depth and have weights (to keep the line vertical), a standard-type zooplankton light trap (commercial or especially designed) every 100 m, and a depth/temperature sensor at 400 and 800 m (to check that the line is kept vertical). Amphipods were attracted by red light as shown by the FishCam on the CTD during the SAS-Oden expedition, but other colours of light could be tested before defining the standard line that will be used within the JPSRM all over the CAO. This will also provide animals for C:N ratios, stable isotopes and fatty acids analyses from specific depths with higher precision than a MIK net.

11. *Design a standard "JPSRM macrozooplankton sampling net"* This should be a closable net so that it is possible to choose a specific depth. During the SAS-Oden expedition a MIK net with a standard diameter of 2 m was used (not closable). This is was large enough to sample macrozooplankton in a representative manner. The usual ring nets

---

[7] Flores H, et al. (on-line manuscript) Sea-ice decline makes zooplankton stay deeper for longer, 09 January 2023, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-2436026/v1]

[8] Ingvaldsen RB, et al. (2023) Under-ice observations by trawls and multi-frequency acoustics in the Central Arctic Ocean reveals abundance and composition of pelagic fauna. Scientific Reports 13:1000 [https://doi.org/10.1038/s41598-023-27957-x]

[9] Van Franeker JA, et al. (2012) The Surface and Under-Ice Trawl (SUIT). Technical Report https://www.researchgate.net/publication/297794282_The_Surface_and_Under_Ice_Trawl_SUIT#fullTextFileContent

(e.g., Nansen net, WP2) are too small, and bigger ones (like the MIK net) do not have a closing mechanism. It should be possible to use a Nansen release mechanism on a large ring net, such as the WP3. Always use standard depths that cover the whole DSL and a bit below, e.g., 1000-500 m, 500-200 m, 200-0 m (to be decided by the JPSRM).

12. *Use a standard "JPSRM mesozooplankton multinet".* For example, the much used "Midi". Decide upon standard depth strata: commonly use in the CAO are the depth intervals 2000-1000-500-200-100-0 m (5 nets). Decide upon standard mesh size – commonly used in the CAO is 150 µm. This net targets mesozooplankton, i.e., the prey of smaller fish such as juvenile polar cod and myctophids.

13. *Use standardised image-based measurements for zooplankton.* The SC07 project results highlight the potential of computer-aided taxonomic analysis using "ZooScan" in combination with deep-learning software as a time-efficient method to quantify zooplankton species composition, abundance and biomass in the context of a monitoring programme such as the JPSRM. The by the SC07 project established training set on "EcoTaxa" will speed up future development of image analyses of macro- and mesozooplankton of the CAO.

14. *eDNA samples should be taken routinely.* Standard methods for eDNA sampling, filter types, extractions, sequencing and bioinformatics should be developed for inter-compatibility of the results. Contamination from humans, and marine fish, squid and shellfish as human food on board or fish bait should be avoided. For bioinformatics analyses the open-source pipelines, including reference databases, designed at SLU can be used. Metagenomic sequencing is preferred since it gives quantitative results.

15. *Use the established stock assessment parameters.*
    Make a standard protocol for the JPSRM, including methodology. For example:
    * Size and age distribution
    * Fulton's condition index (K)
    * gonadosomatic index (GSI)
    * hepatosomatic index (HSI)

16. *Additional recommended parameters to measure from fish samples.*
    Make a standard protocol for the JPSRM, including methodology. For example:
    * Stomach (+ hindgut?) contents
    * C:N ratio in muscle
    * Stable isotopes (δ^{13}C, δ^{15}N) in muscle
    * Fatty acids composition in muscle
    * Otolith δ^{13}C for estimating field metabolic rate
    * Otolith δ^{18}O for temperature reconstructions
    * Standardised methods for population genetics (species-specific)

17. *Deep-sea sediment otoliths* can provide useful data for the JPSRM. Otoliths indicate which species have dominated in a specific area in the past during periods with different environmental conditions, and can also show which species have invaded the area recently. When sampling of living fish is difficult, sediment otoliths can provide an indication of which species to use in stock assessment and modelling. Sampling of deep-sea sediments can be performed during dedicated ecosystem expeditions to the CAO, but also use "ships of opportunity" with no or very limited biological sampling, e.g., geological surveys. Furthermore, the geological research institutions of the CAOFA parties likely host a wealth of sediment samples that could be used to significantly extend the knowledge on past and present fish distributions in the Arctic Ocean.

1. BACKGROUND AND OBJECTIVES

1.1. The CAO Fisheries Agreement (CAOFA)

The EU is a Party to the "Agreement to prevent unregulated High Seas fisheries in the CAO"[10], known as the "CAO Fisheries Agreement" (CAOFA). As such, it will conduct relevant research in the concerned area to contribute to the implementation of the Agreement, in particular, within the Joint Program of Scientific Research and Monitoring (JPSRM). Scientific advice to the steering body, the "Conference of Parties" (COP) of the CAO Fisheries agreement is provided by the "Scientific Coordinating Group" (SCG), the successor of the "Scientific Experts on Fish Stocks in the Central Arctic Ocean" (FiSCAO) and the "Provisional Scientific Coordinating Group" (PSCG), after the Agreement had been ratified by all Parties on 25 June 2021.

1.2. Existing data on the CAO ecosystem

It is widely acknowledged that very little data exists on fish and fish stocks in the Central Arctic Ocean (CAO) and its High Seas area. Considering the fast environmental changes observed in the CAO during the past decades that are expected to increase human activities in the area (including possible fisheries), it is necessary to collect more on-site data to fill important scientific knowledge gaps by collecting new measurements and samples on the CAO ecosystem in the field.

Available data on fish and fish stocks in the CAO have been reviewed by the Scientific Experts on Fish Stocks in the Central Arctic Ocean (FiSCAO) in 2017[11] and in 2018[12], and by the European Fisheries Inventory in the Central Arctic Ocean (EFICA) Consortium in 2020[13]. Recently, in July 2022, the Working Group on Integrated Ecosystem Assessment of the Central Arctic Ocean (WGICA) published for the first time a comprehensive description of the CAO ecosystem[14], which also confirms that more data on the CAO fish stocks are needed. The first scientific paper ever reporting on fish abundance in the central Arctic Deep Scattering Layer (DSL) in the Atlantic water layer (AWL) of the CAO, including the North Pole, was published in 2021[15] and scientists from the EFICA Consortium published the scientific paper "Unexpected fish and squid in the central Arctic deep scattering layer" in the journal "Science Advances" on 18 February 2022[16]. This latter publication is also part of this report.

1.3. Field work in the CAO

The Contracting Authority (CINEA) has on 29 November 2018 contracted the EFICA Consortium to collect new field data that would be necessary to enhance scientific knowledge of the CAO ecosystem, including the existence of possible fish stocks

---

[10] Agreement to prevent unregulated high seas fisheries in the Central Arctic Ocean (Official Journal of the European Union L 73, 15.3.2019, pp. 3-8) [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2019:073:FULL&from=FR]

[11] FiSCAO (2017) Final Report of the Fourth Meeting of Scientific Experts on Fish Stocks in the Central Arctic Ocean, 82 pp. [https://www.fisheries.noaa.gov/event/fourth-meeting-scientific-experts-fish-stocks-central-arctic-ocean]

[12] FiSCAO (2018) Final Report of the Fifth Meeting of Scientific Experts on Fish Stocks in the Central Arctic Ocean, 45 pp. [https://www.fisheries.noaa.gov/event/fifth-meeting-scientific-experts-fish-stocks-central-arctic-ocean]

[13] Snoeijs-Leijonmalm P, et al. (2020) Review of the research knowledge and gaps on fish populations, fisheries and linked ecosystems in the Central Arctic Ocean (CAO). Publications Office of the European Union [https://data.europa.eu/doi/10.2826/387890]

[14] Skjoldal HR (Ed.) (2022) Ecosystem assessment of the Central Arctic Ocean: Description of the ecosystem. ICES Cooperative Research Reports 355:1-341 [https://doi.org/10.17895/ices.pub.20191787]

[15] Snoeijs-Leijonmalm P, et al. (2021) A deep scattering layer under the North Pole pack ice. Progress in Oceanography 194: 102560 [https://doi.org/10.1016/j.pocean.2021.102560]

[16] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]
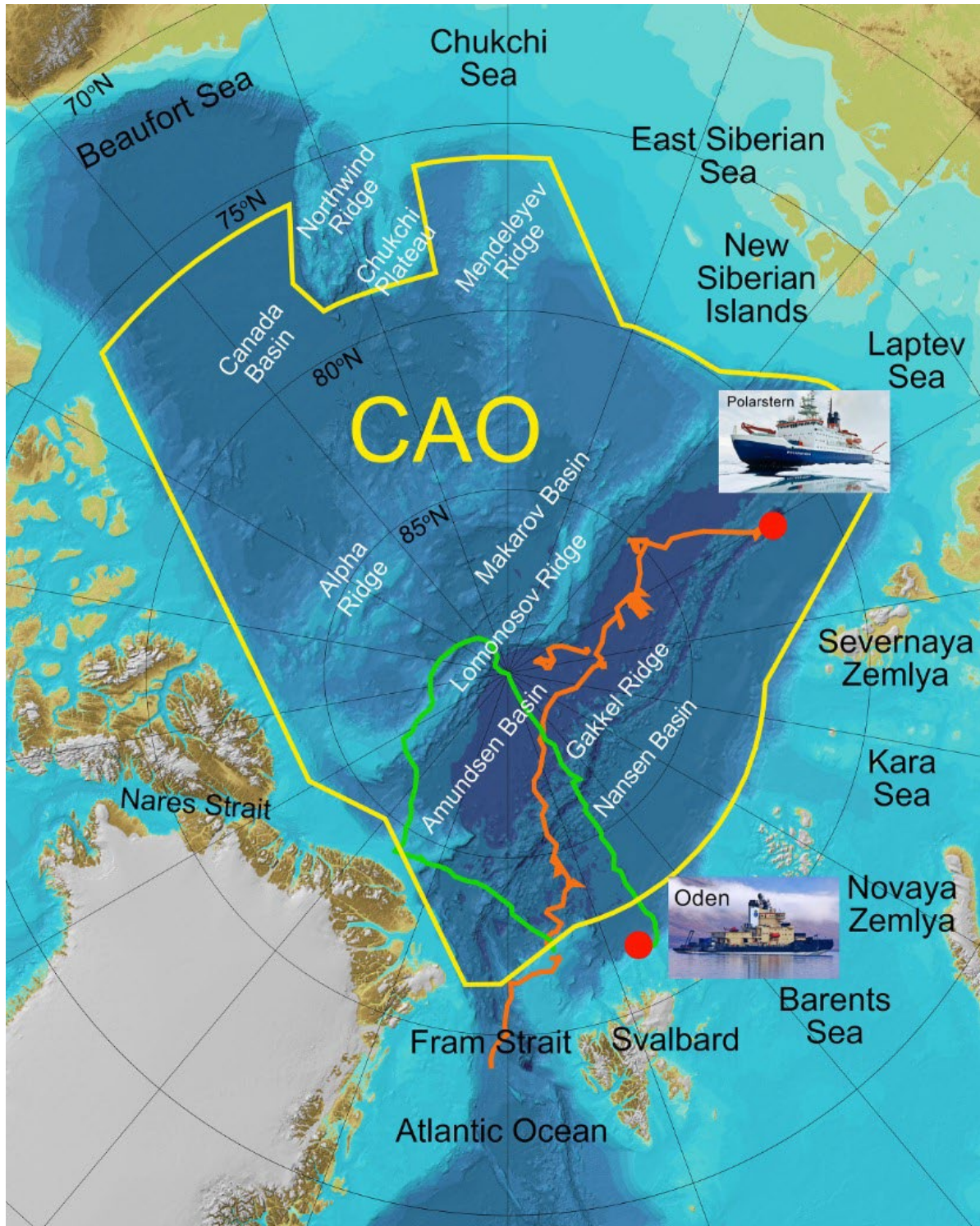
Figure 2: The expedition routes of the MOSAiC and SAS-Oden expeditions, orange = MOSAiC (2019-2020) and green = SAS-Oden (2021). The red dots indicate the start of the expeditions. The MOSAiC expedition consisted of a transpolar drift (Legs 1-3), continued drift south to Fram Strait after a visit to Svalbard for exchanging scientists and crew due to Covid19 (Leg 4), and a final short drift track at 88-89 °N (Leg 5). For the transits (not shown in this map), the only data available are acoustic data, but due to the sound of ice-breaking, most of these data are expected to be disturbed. Special permissions were granted for the SAS-Oden expedition to sample within the EEZs of Greenland and Norway. The background map was extracted from IBCAO[17].

[17] Jakobsson M, et al. (2012) The International Bathymetric Chart of the Arctic Ocean (IBCAO) Version 3.0. Geophysical Research Letters 39:L12609 [http://doi.org/10.1029/2012GL052219]

(Framework Contract EASME/EMFF/2018/003: "Provision of Scientific Support to the High Seas Fisheries in the Central Arctic Ocean"). As part of fulfilling the FWC, the EFICA Consortium has participated in two expeditions to the CAO with the European icebreakers, RV Polarstern in September 2019 - October 2020 and IB Oden in July - September 2021, respectively. The participation of the EFICA Consortium in the expeditions was prepared, a.o., during the EFICA workshop held at the European Commission, DG MARE, Rue Joseph II 99, Brussels on 17 and 18 June 2019[18].

Within SC03[19] of the FWC, EFICA Consortium scientists participated in the international "Multidisciplinary drifting Observatory for the Study of Arctic Climate" (MOSAiC) expedition[20]. The backbone of this expedition was the year-round operation of the German icebreaker RV Polarstern, drifting with the sea ice across the Amundsen Basin, Nansen Basin and Fram Strait to collect ecosystem data reflecting seasonal patterns (**Figure 2**). MOSAiC was the largest Arctic research expedition ever; it included participation of more than 500 scientific experts and 300 support staff from 20 nations. Within SC06[21] of the FWC, EFICA Consortium scientists participated in the "Synoptic Arctic Survey" (SAS)[22] expedition with the Swedish icebreaker Oden[23] (the "SAS-Oden expedition"), steaming for two months from Svalbard to the North Pole, then following the Lomonosov Ridge to the northern Greenland shelf, and then back to Svalbard, collecting ecosystem data reflecting spatial patterns in sea ice conditions during the Arctic summer (**Figure 2**).

These two CAO expeditions provided an excellent opportunity to collect new ecosystem data focusing on fish stock related data at a fraction of the costs for two full expedition budgets. The cost of expeditions to such remote and adverse areas is usually extremely high and therefore had not this opportunity of joining already planned expeditions presented itself, it would have been impossible to fund an expedition solely dedicated to ecosystem and fish stock research. Together, the two expeditions covered a large portion of the Eurasian Basin, the North-Pole area and the western Lomonosov Ridge area (**Figure 2**).

1.4. Objectives of the SC07 project

The objectives of the SC07 project were to elaborate the measurements and samples collected by the EFICA Consortium scientists during the MOSAiC and SAS-Oden expeditions, and to record, store and analyse the results with the aim to obtain conclusions that could contribute to an assessment of the Eurasian part of the CAO ecosystem with respect to potential fish stocks and possible future fisheries in the CAO. Data and sample analysis, recording and storing encompasses results obtained from sample elaboration performed before the signature of the SC07 contract, i.e., the data that were produced by EFICA Consortium scientists in a scientific paper[24].

The analyses performed within the SC07 project concern the measurements and samples covered by the metadata databases delivered to CINEA from the MOSAiC and SAS-Oden expeditions (SC03 and SC06), both from the High Seas area within the CAO and adjacent Exclusive Economic Zones (EEZs) north of Greenland and Norway and in Fram Strait. The SC03 Metadata Database was delivered to CINEA on 17 May 2021 and the SC06

---

[18] EFICA Consortium (2019) EFICA Workshop in preparation for the Polarstern and Oden expeditions for 2019 and 2020. Report to EASME/DG MARE. 32 pp. (EU internal document)

[19] Snoeijs-Leijonmalm P, et al. (2021) Ecosystem mapping in the Central Arctic Ocean (CAO) during the MOSAiC Expedition. Publications Office of the European Union [https://data.europa.eu/doi/10.2926/714618]

[20] https://mosaic-expedition.org/

[21] Snoeijs-Leijonmalm P, et al. (2022) Ecosystem mapping in the Central Arctic Ocean (CAO) during the SAS-Oden Expedition. Publications Office of the European Union [https://data.europa.eu/doi/10.2826/958629]

[22] Synoptic Arctic Survey - a pan-Arctic Research Program [https://synopticarcticsurvey.w.uib.no/files/2019/03/SAS-layout-180629.pdf]

[23] https://www.polar.se/en/expeditions/synoptic-arctic-survey-2021/

[24] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

Metadata Database on 15 January 2022 (**Table 2**). This report occasionally refers to "unpublished data", which are data that were collected during the MOSAiC and SAS-Oden expeditions by non-EFICA scientists or by non-EFICA scientists in collaboration with EFICA. For detailed explanations of the data concerned and how to access these data, see sub-chapters 1.9-1.11 and the first sub-chapters of Chapters 2-7 of this report.

Table 2: Overview of the SC03 (MOSAiC expedition 2019-2020), SC06 (SAS-Oden expedition 2021) Metadata Databases. These databases are available in a _NextCloud_ data storage at Stockholm University, as well as in a _Dropbox_ data storage, that both are accessible for all scientists involved in the SC07 project, as well as for CINEA and DG MARE.

Database	File name
SC03 (MOSAiC)	SC03 Database File 01 - Explanations_210616.xlsx
	SC03 Database File 02 – Logbook Device Operations_210621.xlsx
	SC03 Database File 03 - SHIP_DATA_210616.xlsx
	SC03 Database File 04 - PEL_PHYS_210616.xlsx
	SC03 Database File 05 - ICE_PHYS_210616.xlsx
	SC03 Database File 06 - PEL_BGC_210618.xlsx
	SC03 Database File 07 - ICE_BGC_210618.xlsx
	SC03 Database File 08 - PEL_INV_210621.xlsx
	SC03 Database File 09 - ICE_INV_210618.xlsx
	SC03 Database File 10 - PEL_ACOUST_210621.xlsx
	SC03 Database File 11 - PEL_VIDEO_210618.xlsx
	SC03 Database File 12 - PEL_FISH_210616.xlsx
	SC03 Database File 13 - ICE_FISH_210618.xlsx
	SC03 Database File 14 - PEL_DNA_210621.xlsx
	SC03 Database File 15 - ICE_DNA_210618.xlsx
SC06 (SAS-Oden)	SC06 Database File 01 - Logbook 220115.xlsx
	SC06 Database File 02 - SHIP data 220115.xlsx
	SC06 Database File 03 – SAS CTD metadata 220115.xlsx
	SC06 Database File 04 – EFICA ACOUSTIC metadata 220115.xlsx
	SC06 Database File 05 – EFICA OPTICAL metadata 220115.xlsx
	SC06 Database File 06 – EFICA ZOOPLANKTON metadata 220115.xlsx
	SC06 Database File 07 – EFICA FISH metadata 220115.xlsx
	SC06 Database File 08 – EFICA OTOLITH metadata 220115.xlsx
	SC06 Database File 09 – SAS OMICS metadata 220115.xlsx

1.5. Overarching aim and research questions

The aim of the SC07 project was to analyse data and samples that were previously collected in the CAO for ecosystem mapping and to evaluate how the results could be relevant for future fish stock modelling and assessment in the area. The project can be considered a pilot study for the Joint Program of Scientific Research and Monitoring (JPSRM) that will be carried out in the CAO during the coming 14 years. The project work was subdivided into seven Work Packages (WPs) that executed the tasks as described in the SC07 Terms of Reference and addressed four overarching questions (**Table 3**). Each of the WPs, except WP1 (Coordination), defined a specific set of research questions presented in the respective chapters of this report. The questions overlapped between the WPs which created many positive synergies within the project.

Table 3: The overarching questions to be answered by SC07 and the WPs that addressed these questions indicated in red.

Question in ToR	WP2	WP3	WP4	WP5	WP6	WP7
(1) Which <u>species</u> of nekton (fish and squid) occur in the CAO?		🟥	🟥	🟥	🟥	🟥
(2) What is the <u>distribution, abundance and biomass</u> of nekton and zooplankton?	🟥	🟥	🟥	🟥	🟥	🟥
(3) What is the <u>ecology </u>of the nekton?	🟥			🟥	🟥	🟥
(4) What is the <u>carrying capacity</u> of the CAO with respect to nekton?	🟥	🟥		🟥		

1.6. The organisational structure of the SC07 project

The work in SC07 was carried out jointly by five EFICA Consortium partner institutes (SU, SLU, AWI, WMR, KUL) according to the competence and laboratories available at the respective institutes. Altogether, 22 EFICA Consortium scientists from the partner institutes, performed the sample and data analyses to fulfil the objectives of SC07 (**Table 4, Table 5**).

The SC07 Project Coordinator (Pauline Snoeijs-Leijonmalm, Stockholm University, Sweden) coordinated the work and was responsible for the deliveries to the European Commission as defined in the Inception Report to CINEA. During the Progress, Interim and Final Meetings (**Table 6**), the EFICA Consortium was represented by the Project Coordinator as well as by selected experts who had achieved new results / milestones that should be highlighted and were best presented by the scientists that have achieved these new results themselves.

The tasks of the Project Coordinator of the SC07 project included: (1) daily follow-up of the ongoing work, (2) compilation of the results in collaboration with the scientists responsible for delivering specific results, (3) reporting to CINEA / DG MARE according to **Table 6**. Throughout SC07, the Project Coordinator maintained contacts with CINEA / DG MARE in order to ensure that the tasks were carried out adequately to fulfil the objectives of the SC07 project.

Table 4: The organisational structure of the SC07 project. SIA = stable isotope analysis, FAA = Fatty acids analysis

Work package	Activity	Institute(s)	Task as described in the ToR
WP1	Coordination	SU	All tasks (1-10)
WP2	Hydroacoustics of fish and fish habitats	WMR, SLU, SU	3c,d, 4, 5, 6, 7
WP3	In-situ optics and zooplankton acoustics	AWI, SU, WMR	3c,d, 4, 5, 6, 7
WP4	Fish prey availability and quality (SIA, FAA)	AWI, SU	3c,d, 4, 5, 6, 7
WP5	Fish samples (stomach contents, SIA, FAA)	AWI, KUL, WMR, SU	3c,d, 4, 5, 6, 7
WP6	Sediment otoliths (SIA, ^{14}C dating)	SU, AWI	3c,d, 4, 5, 6, 7
WP7	Environmental DNA (eDNA)	SU, SLU	3c,d, 4, 5, 6, 7

1.7. The scientists contributing to the SC07 project

Altogether, 29 scientists have participated in collecting samples and data in the field during the two expeditions in the CAO (SC03 and SC06) and/or in the elaborations of the samples and data within the SC07 project (**Table 5**). Seven of these scientists were during the SC07 project affiliated with Stockholm University (SU), nine with the Swedish Agricultural University (SLU), five with the Alfred Wegener Institute (AWI), Germany, three with the Catholic University of Leuven (KUL), Belgium, two from Wageningen Marine Research (WMR), the Netherlands, and three with the National Bioinformatics Infrastructure Sweden (NBIS). Seven of the scientists only participated in the field work (SC03 and SC06) while the other 22 participated in the elaborations of the samples and data within the SC07 project.

Table 5: The scientists from the EFICA Consortium that have participated in the field work (SC03, SC06) and/or the sample and data analyses (SC07). * = participation in the SC07 project through an additional grant for method development in bioinformatics from the National Bioinformatics Infrastructure Sweden (NBIS).

Name	Title	EFICA Consortium partner	MOSAiC expedition	SC03	SAS-Oden expedition	SC06	SC07
Pauline Snoeijs Leijonmalm	Professor	SU	MOSAiC	Yes	Oden	Yes	Yes
Serdar Sakinan	Dr	WMR	MOSAiC	Yes	Oden	Yes	Yes
Nicole Hildebrandt	Dr	AWI	MOSAiC	Yes	Oden	Yes	Yes
Anders Svensson	MSc	SLU	MOSAiC	Yes			
Giulia Castellani	Dr	AWI	MOSAiC	Yes			
Hauke Flores	Dr	AWI		Yes	Oden	Yes	Yes
Stefan Bertilsson	Professor	SLU		Yes		Yes	Yes
Moritz Buck	Dr	SLU		Yes		Yes	Yes
Barbara Niehoff	Dr	AWI		Yes		Yes	Yes
Fokje Schaafsma	Dr	WMR		Yes		Yes	Yes
Filip Volckaert	Professor	KUL		Yes		Yes	Yes
Kim Vane	Dr	AWI		Yes			Yes
Jonas Hentati Sundberg	Dr	SLU		Yes			Yes
Julek Chawarski	MSc	SU			Oden	Yes	Yes
Prune Leroy	Dr	SLU			Oden	Yes	Yes
Javier Vargas Calle	Dr	SLU			Oden	Yes	Yes
Flor Vermassen	Dr	SU			Oden	Yes	Yes
Clara Pérez Martínez	MSc	SU			Oden	Yes	
Claudia Morys	Dr	SU			Oden	Yes	
Julia Muchowski	MSc	SU			Oden	Yes	
Baldvin Thorvaldsson	MSc	SLU			Oden	Yes	
Frank Menger	Dr	SU			Oden	Yes	
Hans Nilsson	Dr	SLU					Yes
Sarah Maes	Dr	KUL					Yes
Marie Verheye	MSc	KUL					Yes
Lauren Davies	MSc	SLU					Yes
Allison Churcher *	Dr	NBIS-UMU					Yes
John Sundh *	Dr	NBIS-SU					Yes
Marcin Kierczak *	Dr	NBIS-UU					Yes
Number of scientists		29	5	11	13	18	22

1.8. Time line of the SC07 project

The duration of specific contract SC07 was 16 months in the period January 2022 – May 2023 (**Table 6**). The Gantt chart (**Table 7**) summarises the timing of the analyses and deliveries of SC07.

Each scientist participating in SC07 made a detailed work plan in January 2022 and discussed it with the Project Coordinator with respect to deadlines for delivery (**Table 6**). These work plans were regularly followed up well in time before the Progress and Interim Meetings with CINEA / DG MARE, until the delivery of the final results to the Project Coordinator for inclusion in this Final Report to CINEA. All deadlines were kept by the EFICA Consortium scientists throughout the project.

Table 6: Time table for the deliveries of the SC07 project.

Deliverable	Meetings and Reports	Deadlines	Latest date for submission to CINEA
		T0 = 20 December 2021	T = 16 months (contract ended on 19 April 2023)
1	Draft Inception Report	T0 + 1 week	25 January 2022
2	Kick-off Meeting Discussion Draft Inception Report Discussion implementation requirements Discussion expected deliveries	T0 + 2 weeks	2 February 2022
3	Inception Report	T0 + 1 month	14 February 2022
4	Progress Meeting PowerPoint presentation	T0 + 4.5 months	6 May 2022
5	Progress Meeting Discussion on the progress of the analyses Discussion achieved results	T0 + 5 months	20 May 2022
6	Draft Interim Report	T0 + 8 months	5 September 2022
7	Interim Meeting Discussion on the progress of the analyses Discussion achieved results	T0 + 8.5 months	16 September 2022
8	Interim Report	T0 + 9 months	30 September 2022
9	Draft Final Report	T0 + 14 months	20 February 2023
10	Final meeting Discussion of all results Wider audience, incl. stakeholders	T0 + 14.5 months	3 March 2023
11	Final Report Including Database and Manual	T0 + 16 months	19 April 2023

1.9. Three types of data

This report contains the results for three types of data as summarised in Table 9 of the SC07 Inception Report and the MOSAiC and SAS-Oden Data Policies (included as appendices to this report). The data type is especially important with respect to data ownership and conditions for publication.

The first type of data are expedition raw data owned by the Alfred Wegener Institute (AWI: MOSAiC ship route data, ship-based acoustic data, CTD data) or the Swedish Polar Research Secretariat (SPRS: SAS-oden ship route data, ship-based acoustic data, CTD data), or raw DNA and RNA sequencing data owned collectively by the scientists who collected the data. All these data are published with open access, but if they are to be used for publications, the contact person as indicated in the database should be contacted to discuss the purpose of the publication. Often this results in a collaboration regarding publication. These data are already available in public databases when this report is delivered to CINEA or will be soon after (within 2023).

Table 7: Gantt Chart of the SC07 project.

Task	Responsible	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M
Work Package 1: Organisation																		
Cloud-based account	Project Coordinator	■																
Update sample lists	Project Coordinator	■																
Synchronization ICES	Project Coordinator	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Draft Inception Report	Project Coordinator	■																
Inception Meeting	Project Coordinator	■																
Inception Report	Project Coordinator			■														
Progress Meeting	Project Coordinator						■											
Progress PowerPoint	Project Coordinator						■											
Draft Interim Report	Project Coordinator									■								
Interim Meeting	Project Coordinator									■								
Interim Report	Project Coordinator									■								
Draft Final Report	Project Coordinator															■		
Final Meeting	Project Coordinator															■		
Final Report	Project Coordinator																	■
Work Package 2: Hydroacoustics of fish and fish habitats																		
WP2 analyses	WMR, SLU, SU	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	
Work Package 3: In-situ optics and zooplankton acoustics																		
WP3 analyses	AWI, WMR, SU	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	
Work Package 4: Fish prey availability and quality																		
WP4 analyses	AWI, SU	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	
Work Package 5: Fish samples																		
WP5 analyses	AWI, WMR, KUL, SU	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	
Work Package 6: Sediment otoliths																		
WP6 analyses	AWI, SU	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	
Work Package 7: eDNA																		
WP7 analyses	SLU, SU	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	▒	

Months (January 2022 – May 2023)

The second type of data are SC07-specific data produced by the EFICA Consortium exclusively within the SC07 project. These data are owned by the EU Commission. If these data are to be used for publications, written permission must be requested from CINEA / DG MARE. The EFICA Consortium is planning for peer-reviewed scientific publication using these data in combination with data provided by other scientists. The EFICA Consortium will through its Coordinator Pauline Snoeijs-Leijonmalm request written permission from CINEA / DG MARE when preliminary manuscripts are available even after the FWC has ended. Most preliminary manuscripts (partly) containing this type of data are expected to be ready within half a year after the end of the SC07 project.

The third type of data are data produced by the EFICA Consortium within the SC07 project in collaboration with other scientists participating in the MOSAiC and SAS-Oden expeditions or data belonging entirely to other scientists but that could be of use for modelling and assessment (e.g., productivity data) in the future. Collecting these data by the EFICA Consortium scientists alone would not have been possible, both practically because the field work was necessarily highly collaborative in harsh conditions, and also with respect to other scientist's projects being financed by their own national research councils. These data are published on line with open access after a peer-reviewed scientific publication has been published. Most preliminary manuscripts containing this type of data are expected within half a year after the end of the SC07 project and will be shared with CINEA / DG MARE.

1.10. Data storage of the SC07-specific data

Throughout the project (**Table 7**) it was possible for all involved in the SC07 project, including CINEA and DG MARE, to follow the progress of the SC07 project in the EFICA

NextCloud and *Dropbox* data storages with the same contents in both storages. Other information relevant for the implementation of the SC07 project was stored here as well. At the end of the SC07 project the data files with the final results (**Table 8**) were made available in both depositories. The expiration date of these accounts is no earlier than 31 December 2023. No password is required, but only people with granted access can enter via one of the following links:

https://su.drive.sunet.se/index.php/apps/files/?dir=/200Gb_userdata/EFICA%20Consortium&fileid=3287558

https://su.drive.sunet.se/index.php/s/2Crsd4Lj9M4bf2s

https://www.dropbox.com/home/Data%20delivery%20to%20CINEA

Table 8: Elaborated data (project results) in the EFICA SC07 database delivered with this SC07 Final Report. These data files are available in a _NextCloud_ data storage at Stockholm University, as well as in a _Dropbox_ data storage, that both are accessible for all scientists involved in SC07, as well as for CINEA and DG MARE.

Work package	Work package name	Data file name
WP1	EFICA SC07 database manual	EFICA_DATA_SC07-WP1.pdf
WP2	Hydroacoustics of fish and fish habitats	EFICA_DATA_SC07-WP2_A.variables.xlsx EFICA_DATA_SC07-WP2_B.EK80_MOSAIC.csv EFICA_DATA_SC07-WP2_C.EK80_SAS-Oden.csv EFICA_DATA_SC07-WP2_D.WBAT333KHz.csv
WP3	In-situ optics and zooplankton acoustics	EFICA_DATA_SC07-WP3_A.UVP.xlsx EFICA_DATA_SC07-WP3_B.LOKI.xlsx
WP4	Fish prey availability and quality	EFICA_DATA_SC07-WP4.xlsx
WP5	Fish samples	EFICA_DATA_SC07-WP5.xlsx
WP6	Sediment otoliths	EFICA_DATA_SC07-WP6.xlsx
WP7	Environmental DNA (eDNA)	EFICA_DATA_SC07-WP7.xlsx

1.11. Data storage of other relevant data

Both expeditions have their home database where all metadata and data are expected to be published according to the MOSAiC and SAS-Oden data policies (included as appendices to this report). Nobody was allowed to participate in the ship-borne field work without signing the expedition data policy. For the MOSAiC expedition in 2019-2020 the home database is *Pangaea* (hosted by AWI)[25] and for the SAS-Oden expedition in 2021 the home database is the *Bolin Centre Database* (hosted by SU)[26]. During the coming years these depositories will continuously be populated with an increasing number of data files with expedition results.

Publication of the acoustic raw data is as follows:
MOSAiC Leg 1: https://doi.pangaea.de/10.1594/PANGAEA.916105
MOSAiC Leg 2: https://doi.pangaea.de/10.1594/PANGAEA.920008
MOSAiC Leg 3: https://doi.pangaea.de/10.1594/PANGAEA.923587
MOSAiC Leg 4: Data publication (Pangaea) soon, responsible is Sören Krägefski (AWI)
MOSAiC Leg 5: Data publication (Pangaea) soon, responsible is Sören Krägefski (AWI)
SAS-Oden: Data publication (Bolin) soon, responsible is Julia Muchowski (SU)

---

[25] PANGAEA Data Publisher for Earth & Environmental Science [https://www.pangaea.de]
[26] Bolin Centre Database, Stockholm University [https://bolin.su.se/data/]

Publication of the CTD data is as follows:
MOSAiC: https://doi.org/10.1594/PANGAEA.936275
(this is a preliminary dataset, responsible for publication is Sandra Tippenhauer (AWI)
SAS-Oden: https://doi.org/10.1594/PANGAEA.951264 (already published)

Publication of raw sequencing data is as follows:
MOSAiC: Doe Joint Genomics Institute, USA: https://img.jgi.doe.gov/cgi-bin/mer/main.cgi
Example:
https://img.jgi.doe.gov/cgi-bin/mer/main.cgi?section=TaxonDetail&page=taxonDetail&taxon_oid=3300060642
Download guide:
https://groups.google.com/a/lbl.gov/forum/#!msg/img-user-forum/o4Pjc_GV1js/EazHPcCk1hoJ
SAS-Oden: EMBL's European Bioinformatics Institute (https://www.ebi.ac.uk)
About 1000 metagenomes and metatranscriptomes will be uploaded during spring 2023

1.12. Update of available data and samples

In accordance with the SC07 project Terms of Reference, the Project Coordinator has, in collaboration with the responsible scientists, updated the lists of collected field measurements and samples and decided which samples to analyse. A unique ID number are provided for each sample. All samples taken within the SC03 and SC06 projects (**Table 2**), but not all replicates, taken during both expeditions were analysed to achieve the objectives of SC07 as described in the Inception Report. Out of precaution, replicates were taken if possible to have back-ups in case samples would get damaged and for archiving (a standard procedure for, e.g., DNA and RNA). Some analyses are very expensive and no extra information is obtained by analysing more replicates than planned compared to the analysis costs (e.g., fatty acids). For appropriate parameters, test series were made to determine at which number of replicates standard deviations of the mean or multivariate community analysis stabilise. In this report the number of samples analysed within SC07 are specified in the respective chapters.

1.13. Identification of already analysed data and samples

In accordance with the SC07 project Terms of Reference, the Project Coordinator identified the data obtained from field measurements and sample elaboration performed before the signature of the SC07 Contract, i.e., the data that were produced by the EFICA Consortium scientists for the scientific paper "Unexpected fish and squid in the central Arctic deep scattering layer"[27], and the field measurements and samples that have not been analysed before the signature of the specific contract SC07. The data on which the international publication is based have been deposited in the EFICA *NextCloud* and *Dropbox* depositories and are also available on-line at the Science Advances journal site with the published paper. This paper is freely accessible on-line ("open access").

---

[27] Snoeijs-Leijonmalm P. et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

2. HYDROACOUSTICS OF FISH AND FISH HABITATS (WP2)

2.1. Research questions addressed by WP2

(1) What is the potential fish abundance along the MOSAiC and SAS-Oden expedition routes as estimated from acoustic data?
(2) What is the potential fish biomass along the expedition routes of the MOSAiC and SAS-Oden expeditions as estimated from acoustic data?
(3) How does fish abundance and biomass relate to zooplankton abundance and biomass? (collaboration with WP3)
(4) What is the physical and biotic environment in which the potential fish resided during the two expeditions?

2.2. Data produced by WP2

EFICA_DATA_SC07-WP2_A.variables.xlsx
EFICA_DATA_SC07-WP2_B.EK80_MOSAIC.csv
EFICA_DATA_SC07-WP2_C.EK80_SAS-Oden.csv
EFICA_DATA_SC07-WP2_D.WBAT333KHz.csv

For details of the Device Operations (date, time, geographical position, station depth), see files "MOSAiC_Device_operations" and "SAS-Oden_2021_Logbook"[28].

The raw CTD data are publicly available via PANGAEA[29] and the Bolin Centre Database[30] for MOSAiC and SAS-Oden, respectively. They are owned by the AWI and the SPRS but they cannot be used in publications without approval of the scientists who collected these data because the oceanographers have not published their scientific papers yet. It is expected that these publications come out during 2023.

2.3. Human resources of WP2 and main responsibilities

Serdar Sakinan (WMR) coordination of the acoustic data analyses, advanced data processing; Hans Nilsson and Jonas Hentati-Sundberg (SLU) initial data screening; Christian Stranne (SU) physical oceanography

2.4. Methods used by WP2

Short description of the available data sets

The acoustic data consisted of two main data sets: (1) MOSAiC data collected on board of the German icebreaker RV Polarstern during the international MOSAiC drift expedition 2019-2020, and (2) SAS-Oden data collected on board of the Swedish icebreaker Oden during the Swedish Synoptic Arctic Survey (SAS) in 2021. The MOSAiC data consist of continuous echosounder recordings covering one year period throughout the transpolar ice drift. The SAS-Oden data set similarly consisted of echosounder recordings as well as vertical profiles acquired by an autonomous echosounder (WBAT) that can be lowered to depths down to 1000 m when the ship is stationary. Details of data collection are available in two previous reports[31,32].

---

[28] Bolin Centre Database [https://bolin.su.se/data/]
[29] PANGAEA Data Publisher for Earth & Environmental Science [https://www.pangaea.de]
[30] Bolin Centre Database [https://bolin.su.se/data/]
[31] Snoeijs-Leijonmalm P, et al. (2021) Ecosystem mapping in the Central Arctic Ocean (CAO) during the MOSAiC Expedition. Publications Office of the European Union [https://data.europa.eu/doi/10.2926/714618]
[32] Snoeijs-Leijonmalm P, et al. (2022) Ecosystem mapping in the Central Arctic Ocean (CAO) during the SAS-Oden Expedition. Publications Office of the European Union [https://data.europa.eu/doi/10.2826/958629]

MOSAiC: data collection and calibrations

Acoustic data were continuously collected during the transpolar drift between 28 September 2019 and 2 October 2020 using a Simrad EK60/80 echosounder with hull-mounted 38, 70, 120 and 200 kHz transducers. The data were processed with main focus on fish-like targets with an aim to estimate the density distribution. The density of the deep scattering layer (DSL) varied along the MOSAiC drift track from a few fish-like individuals to thousands within the beam.

A series of calibration tests were carried out on 17 April and 5 May onboard RV Polarstern. The standard calibration sphere was lowered through the moon pool (an opening in the base of the hull allowing access to the water) and the spheres was operated with a cabled remotely operated underwater vehicle (ROV). Based on initial analysis using the Simrad calibration module, the tests were deemed successful for all frequencies, 38, 70, 120 and 200 kHz in CW mode and 70 and 120 kHz echosounders in FM mode. Further analyses were performed using the "Calibration Assistant Function" in the data program Echoview™ and the results are presented in **Table 9** and **Table 10**. Part of these calibrations were used for the acoustic analyses of zooplankton (see WP3).

Table 9: Calibration results of the narrowband signals.

Transducer	ES38B	ES70-7C	ES120-7C	ES200-7C
Frequency	38000	70000	120000	200000
Gain	24.32	23.34	25.99	24.43
SaCorrection	-0.43	-0.07	-0.16	-0.34
BeamWidthAlongship	6.86	8.72	5.99	6.93
BeamWidthAthwartship	6.76	7.03	5.89	6.94
TsRmsError	0.0518	0.6421	0.0898	0.0691
Pulse Length	1.024	1.024	0.256	1.024
No of hits	1751	861	1540	1356

MOSAiC: data screening and cleaning
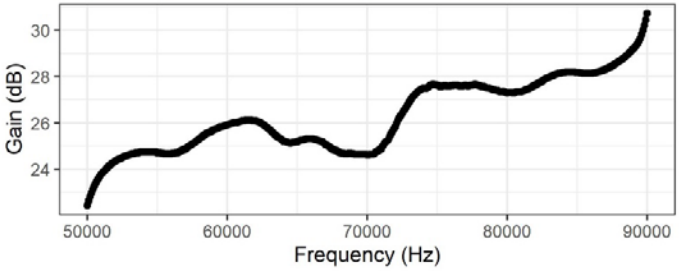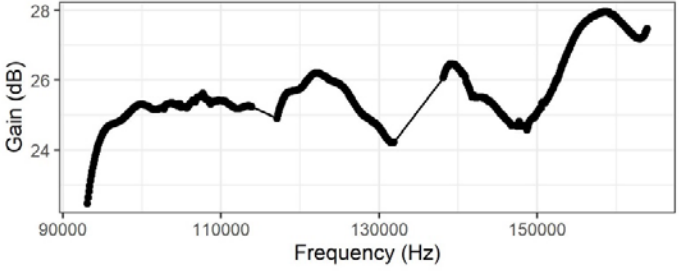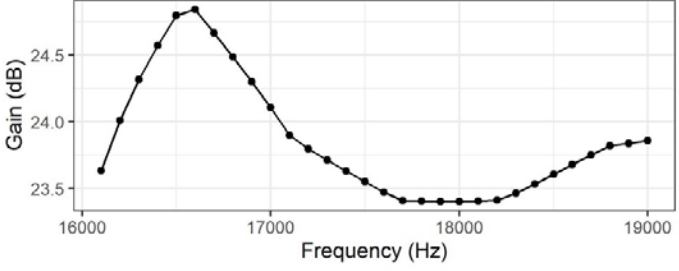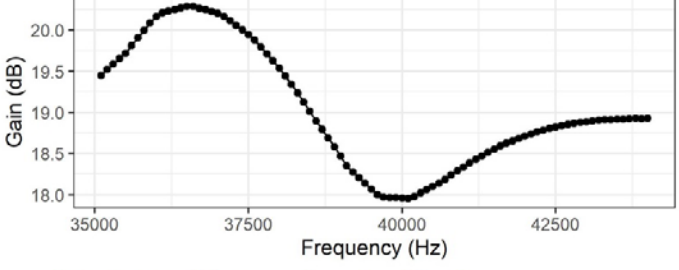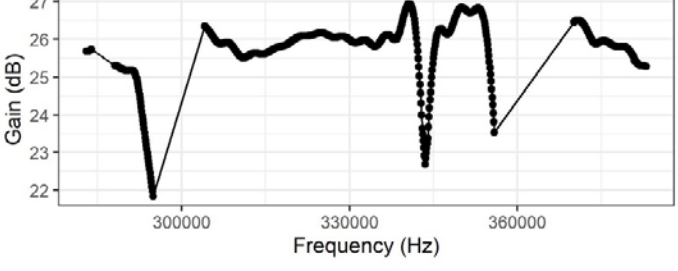
Raw acoustic raw data contains unusable parts such as excessive noise and reflections from unintended targets. Noise becomes a problem when it suppresses the desired signals from the targeted organisms. Such noise is usually generated by the ship itself when the engines/propellers are running, or by other acoustic instruments (e.g., Multibeam, ROV, ADCP), electrical noise due to interference from the vessel's power system, or other electromagnetic interference somehow picked up by the echosounder system. In general, it was possible to visually identify such noise on the echograms, assess their level of significance and remove from the data set.

Other unwanted parts of the raw data consisted of the backscatter from unintended targets, such as equipment lowered under water (e.g., CTD rosette, zooplankton nets, ship maintenance equipment), air bubbles and sea-ice particles pushed downwards during steaming, echoes from the seabed, or secondary bottom signals ("false bottom") especially in regions close to the continental shelf. Before exporting any quantitative metrics from the acoustic data, such unusable parts were identified and cleaned. Echograms (visual representations of acoustic data) were generated and visually scrutinised for disturbances.

In addition, a preliminary echo integration was performed to analyse regions of irregular data. Signal-to-noise ratio (SNR) was used as quality metric was with a threshold limit of 10 dB (e.g., when the minimum expected backscatter from fish is -70 dB, the maximum noise could be -80 dB). The sections where noise was stronger than this threshold, these sections were considered as no-data sections. Part of the noise was tackled automatically using the "Background Noise Removal", "Impulse Noise Removal" and "Transient Noise

Removal" tools in Echoview™ in multiple steps separately for each frequency. For spike-noise, mainly "Impulse Noise Removal" was used.

Table 10: Calibration results of the broadband signals.

Equipment	Signals
RV Polarstern 70 kHz Broadband calibration parameters from the test conducted during the MOSAiC expedition. The values in the plot are applied to raw data for the analysis as final calibration gain for the 50-90 kHz bandwidth.	Polarstern 70 kHz Broadband Calibration
RV Polarstern 120 kHz Broadband calibration parameters from the test conducted during the MOSAiC expedition. The values in the plot are applied to raw data for the analysis as final calibration gain for the 95-165 kHz bandwidth.	Polarstern 120 kHz Broadband Calibration
IB Oden 18 kHz Broadband calibration parameters from the test conducted prior to the SAS-Oden expedition. The values in the plot are applied to raw data for the analysis as final calibration gain for the 16-19 kHz bandwidth.	ODEN 18 kHz Broadband Calibration
WBAT 38 kHz Broadband calibration parameters from the test conducted during to the SAS-Oden expedition on ice. The values in the plot are applied to raw data for the analysis as final calibration gain for the 35-44 kHz bandwidth.	WBAT 38 kHz Broadband Calibration
WBAT 333 kHz Broadband calibration parameters from the test conducted during to the SAS-Oden expedition on ice. The values in the plot are applied to raw data for the analysis as final calibration gain for the 283-383 kHz bandwidth.	WBAT 333 kHz Broadband Calibration

Near-surface disturbances were cleaned using a line initially placed at 7 m below the transducers and manually edited depending on the extent and intensity of the disturbances. The backscatter from unintended targets were masked by manually drawn regions and removed from the analysis. The periods when the ship was steaming or breaking ice contained the most significant noise. When the ship was stationary, the main disturbances originated from the grey water hose that was lowered down to 150 m several times per week during the MOSAiC expedition. Seabed echoes, false bottom echoes and surface reverberations (e.g., bubbles, transducer ringdown effects) were also removed manually.

MOSAiC: identification of single targets using and target tracking

The reference frequency for the data analysis targeting fish was 38 kHz. This frequency has been in use as a standard for hydroacoustic studies on fish distributions for many decades and is widely accepted by the scientific community. Other frequencies are often assessed relative to it for identification as there can be slightly different returns at different frequencies from the same target (WGIPS 2022). The 38 kHz system has the most powerful transducer of the multi-frequency set and has the widest range of coverage due to the strength of the signals and relatively little absorption. The assessment of the data showed 600 meters as the deepest depth for detecting reliable tracks from individual fish, and this depth was taken as the lower limit for the analysis. Volume backscatter data also indicated that the backscatter attributable to fish is negligible below this depth, hence 600 m is the lower limit for the data analyses.

During the MOSAiC expedition, the 70 kHz data was collected in broadband and narrowband mode in daily alternation. As the broadband data allows for pulse compression, it is powerful in precise detection of single targets. Therefore, the broadband data was also processed for target tracking as a cross-check for the 38 kHz detections. The 120 kHz data was largely noisy and had relatively limited range, therefore it was deemed not useful density estimation of fish. 200 kHz also das limited range and used mainly for zooplankton analysis and interpretation of the targets detected by 38 kHz within 200 m. The acoustic data analysis was carried out in line with previously published data[33] where the two types of results were generated: (a) "Full Data Set" and (b) "Target-tracking Data Set".

MOSAiC: Full Data Set

The Full Data Set consists of NASC in m^2 nmi^{-2}, which is a result of echo integration using thresholds of -82 dB, -75 dB, and -70 dB. This is the bulk volume backscatter where the individual strength of the targets cannot be discerned (i.e., one large target will produce the same result as multiple small targets). The advantage of the Full Data Set is that backscatter from the organisms will still be taken into account even though they cannot be detected as individual targets, e.g., when fish form dense aggregations or schools. Another advantage is that it can be applied with more consistency to the entire data set and allows uninterrupted time series for spatio-temporal mapping/comparisons. The disadvantage of the Full Data Set is that it could be more susceptible to noise.
MOSAiC: Target-tracking Data Set

The Target-tracking Data Set consists of the detected individual targets that have been successively ensonified multiple times so that they can be detected as fish tracks by the Echoview "Tracking Algorithm". For quantification of fish abundance, target-tracking gives more accurate results when all individual tracks are distinct and isolated, which is nearly always the case in the CAO. The Target-tracking Data Set can be more informative than the Full Data Set because it contains multiple hits from individuals that

---

[33] Snoeijs-Leijonmalm P. et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

allow for statistical estimates such as variability in target strength (TS). Disadvantage of the Target-tracking Data Set occurs when the targets are too close to each other (e.g., high-density aggregations) and the individuals cannot be discerned. It should also be taken into account that the susceptibility of overlap between different targets increases with depth as the acoustic beam widens when it travels further downwards. This may cause bias in high-density regions and deeper portions of the water column. Although this limits the use of fish tracking / echo counting for overall abundance estimation, the method can be especially useful for feeding the echo integration (i.e., estimating the mean TS) and allow verification of density estimates.

MOSAiC: processing target tracks for estimation of target strength (TS)

Tracks are basically a series of successively detected single targets that represent the same fish individual. Target-tracking is performed over the valid sections of the entire echogram were the targets meet the requirements. In order to be accepted as valid, a detected track had to contain at least five valid pixels and the track should last at least for seven pings, even if there are empty pixels. Further criteria for accepting detections were the distances between the observations in three dimensions. These 3D parameters make use of GPS recordings, the vessel's motion sensor (e.g., heave), depth estimation by the echosounder, and angular positional information estimated by the split-beam functionality. The selected dimensions are being sideway, transverse and vertical with maximum allowable distances of 6 m, 6 m and 1.1 m, respectively for successive pings. These parameters were selected empirically after several trial and error procedures using different situations.

To convert the echo integration into the number of individuals, the mean TS of the individuals is needed for the integrated volume. The TS data can be obtained from in-situ measurements, but individual in-situ detections of TS are highly variable due to stochasticity. The average of the multiple hits from the same individual provides a more reliable value and allows for filtering erratic detections, e.g., based on the standard deviation (SD). The mean values can be used to convert the echo integration values into number of individuals. For this, the echograms were gridded in 20-m vertical and 1-h horizontal bins. Based on these grids, both the fish tracks and the echo integration values were extracted. If a cell has an echo integration value but misses the target strength, the value of the nearby cell is taken by gradually increasing the window side horizontally and vertically.

While tracked targets are generated based on individual pixels from the "single target echograms" that are assigned to an individual fish with its true location (i.e., exact 3D position within the beam based on detected angle relative to the centre of the beam) their representation in the volume backscattering echogram can be slightly different as the exact edge of the echo-trace could be fuzzy spreading over larger number of pixels. Furthermore, there could be gaps in the fish tracks where some of the data points are rejected (false negative) by the single target detection operator even when they belong to the same target. In order to compensate for such gaps, an edge enhancement step was introduced where, each fish track region (after the automatic detection) was used as a boolean mask (i.e., inside region is "True" outside is "False" and a dilation filter of 5´5 is applied over these masks to contain the volume backscatter values from the fuzzy edges as dilation process thickens the edges. The advantage of this procedure is that it improves the precision of the measurements in the low density zones where all fish-related echoes are contained in the extracted data set while all non-fish values are excluded.

MOSAiC: estimation of abundance and biomass

In summary, the Target-tracking Data Set provides estimates of the minimum potential number of fish while the Full Data Set gives a maximum estimate of the areal densities of

scatterers[34]. In an ideal case, abundance and biomass estimates from the acoustic backscatter rely on representative samples from the fish community under study. Due to the absence in such representative samples in this study, estimates based on two scenarios were used. The main factor that determines the TS is target size. Taxon-specific target strength is typically expressed with its length dependency as TS = m*log10(L)-b. Where m is the slope and b is the intercept of this regression model. For the analyses TS/length relationship from the literature were used. Sporadic capture of some specimens with longlines and nets during this study (see WP5), in combination with underwater video recordings (see WP3), provided some indication of the possible presence of the species and enabled relatively realistic assumptions.

Candidate species are glacier lanternfish, polar cod *Boreogadus saida*, ice cod *Arctogadus glacialis*, and other North Atlantic species that were captured during the MOSAiC drift: Atlantic cod *Gadus morhua*, haddock *Melanogrammus aeglefinus*, and beaked redfish *Sebastes mentella*. Final estimates were made based on assumptions, taking into account potential taxonomic groups and size distributions.

As the individual TS distribution was obtained from the target-tracking analysis, this information can be used to estimate the length distribution of the targets based on hypothetical assumptions using any available information from the MOSAiC samples (e.g., longline catches) and the target strength equations in **Table 11**. Similarly, the biomass can then be estimated using the length/weight relationship for these species.

Table 11: Target strength using body length relationships from the literature.

Species	m (slope)	b (intercept)	Region	Reference
Gadoids	21.3	-68.3	Lofoten, Norway	Foote (1987)[35]
Gadoids	20	-66.3	Lofoten, Norway	Foote (1987)
Haddock	20	−67.9	Varanger, Norway	Ona & Hansen (1986)[36]
Physoclisti	20	-67.5	Bering Sea, Norwegian Sea	Foote (1987)
Atlantic Cod	20	-66	Newfoundland	Rose & Porter (1996)[37]
Beaked redfish	20	-68.7	Newfoundland	Gauthier & Rose (2002)[38]
Capelin	20	−70.3	North Pacific	Guttormsen & Wilson (2009)[39]
Polar cod	8.03	-60.78	Bering Sea	Parker-Stetter et al. (2011)[40]
Polar cod	14.33	-65.13	Beaufort Sea	Geoffroy et al. (2016)[41]
Polar cod	21.8	-72.7	North Atlantic	Snoeijs-Leijonmalm et al. (2021)[42]

[34] Snoeijs-Leijonmalm P. et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

[35] Foote KG (1987) Fish target strengths for use in echo integrator surveys. The Journal of the Acoustical Society of America 82:981 [https://doi.org/10.1121/1.395298]

[36] Ona E & Hansen K (1986) In-situ target strength observations on haddock. ICES CM 1986/B:39 [https://imr.brage.unit.no/imr-xmlui/handle/11250/104127]

[37] Rose GA & Porter DR (1996) Target-strength studies on Atlantic cod (*Gadus morhua*) in Newfoundland waters. ICES Journal of Marine Science 53:259-265 [https://doi.org/10.1006/jmsc.1996.0032]

[38] Gauthier S & Rose GA (2002) In situ target strength studies on Atlantic redfish (*Sebastes* spp.). ICES Journal of Marine Science 59:805-815 [https://doi.org/10.1006/jmsc.2002.1248]

[39] Guttormsen MA & Wilson CD (2009) In situ measurements of capelin (*Mallotus villosus*) target strength in the North Pacific Ocean. ICES Journal of Marine Science 66:258-263 [https://doi.org/10.1093/icesjms/fsn205]

[40] Parker-Stetter SL, et al. (2011) Distribution of polar cod and age-0 fish in the US Beaufort Sea. Polar Biology 34:1543-1557 [https://doi.org/10.1007/s00300-011-1014-1]

[41] Geoffroy M, et al. (2016) Vertical segregation of age-0 and age-1+ polar cod (*Boreogadus saida*) over the annual cycle in the Canadian Beaufort Sea. Polar Biology 39:1023-1037 [https://link.springer.com/article/10.1007/s00300-015-1811-z]

[42] Snoeijs-Leijonmalm P, et al. (2021) A deep scattering layer under the North Pole pack ice. Progress in Oceanography, 194:102560 [https://doi.org/10.1016/j.pocean.2021.102560]

SAS-Oden: processing of the EK80 data

EK80 echosounder data were collected during the entire SAS-Oden expedition focusing on the upper portion of the water column (20-600 m). The split-beam Simrad ES18-11 transducer of IB Oden was coupled to an EK80 transceiver and used in wideband mode to generate frequency modulated signals. Based on previous experiences with this special configuration, initial tests were carried out prior to the onset of expedition transects as it was known that the signal quality deteriorates towards the higher end of the full bandwidth. While the full bandwidth that can be achieved with this transducer extends from 15 to 28 kHz, the test results showed that the cleanest data for observation of biological targets would be achieved by using a relatively narrow bandwidth and the following pulse parameters: Frequency Start = 16 kHz; Frequency End = 19 kHz; Pulse Duration = 4.096 milliseconds; Sample Interval = 0.256 milliseconds; Transmit Power = 1600 Watt.

Since the density of biological targets along the SAS-Oden expedition route was very low, the focus was on detecting the targets individually, which would allow a detailed understanding of behaviour and size distribution. This required receival of multiple hits from the same target which then would allow for isolation and tracking of the target. As it was necessary to operate the echosounder with short intervals between the pings, the observation range was limited to 1200 m of depth. The ping rate was manually adjusted to keep the aliasing signals from the bottom returns (false bottom) out of the observed range. This adjustment generally allowed for a ping rate of 3 seconds. From a general overview of all data from the SAS-Oden expedition, it was estimated that the usable portion of the data set is ca. 42%. Icebreaking was the major disturbance and generated almost no usable data. When the engines were turned off and the ship was drifting with the ice there were other sources of noise. A regular noise that significantly disturbed data collection was that from a steam hammer in the ship's fuel heating system. This noise was in the form of irregular stripes becoming stronger with depth that were especially amplified by the time-varied gain. There was also a reoccurring temporary elevated broadband background noise that was potentially caused by power usage. Although continuous communication was sustained with the ship's crew, and generally ad-hoc solutions were found for temporary improvements, noise remained to be a perturbation of the data quality. As a result, some portions of the data have substantial noise similar to previous acoustic studies with IB Oden[43].

The data processing for the SAS-Oden EK80 data was performed with the same methods as described for the MOSAiC data. As the noise was heavier in this data set compared to MOSAiC, noise cleaning involved more extensive manual processing. The steaming/ice-breaking portions were unusable for quantitative analyses. Therefore, focus was on the data collected when the ship was stationary. In these relatively cleaner sections, the main sources of noise were the ship's steam hammer, equipment lowered under water such as CTD rosette, zooplankton nets, LOKI, and echoes from the seabed or secondary bottom signals ("false bottom") especially in regions close to the continental shelf. Before exporting any quantitative metrics from the acoustic data, unusable parts were identified and cleaned. Similar to the MOSAiC data, signal-to-noise ratio (SNR) was used as quality metric with a threshold limit of 10 dB. The sections where noise was stronger than this threshold were considered as no-data sections. Some of the noise could be tackled automatically using the "Background Noise Removal", "Impulse Noise Removal" and "Transient Noise Removal" tools in Echoview™ in multiple steps separately for each frequency. or spike-noise, mainly "Impulse Noise Removal" was used. For example, in a few sections where the Multibeam and EK80 needed to be operated together, interference spikes were formed, which were handled using spike filters with further manual processing. Also, it was possible to clean some of the steam hammer noise with the "Transient Noise Removal" tool, but it was necessary to finalise the processing with manual data editing (**Figure 3**).

---

[43] Snoeijs-Leijonmalm P, et al. (2021) A deep scattering layer under the North Pole pack ice. Progress in Oceanography, 194:102560 [https://doi.org/10.1016/j.pocean.2021.102560]

Because of the significance of the noise problem, it was not possible to use the "Full Data Set" for SAS-Oden as for MOSAiC. The focus was mainly on the clean SAS-Oden sections where fish tracks could be identified. If a region contained echo-traces that could be attributed to fish but not captured by the tracking algorithm, a manual box was drawn around that section to include these sections together with the tracks. After defining all fish-containing regions in the data, one last check was performed by using an echo-integration curve, and carefully inspecting the entire echogram for any unnoticed noise.

Near-surface disturbances were cleaned using a line initially placed at 20 m below the transducer and manually editing the data depending on the extent and intensity of the disturbances. The fact that this offset was larger than for MOSAiC was due to a larger transducer ringdown effect in the 18 kHz broadband setup. The backscatter from unintended targets were masked by manually drawn regions and removed from the analysis. Seabed echoes and false bottom echoes and surface reverberations (e.g., bubbles, transducer ringdown effects) were also removed by manually drawn regions and lines.
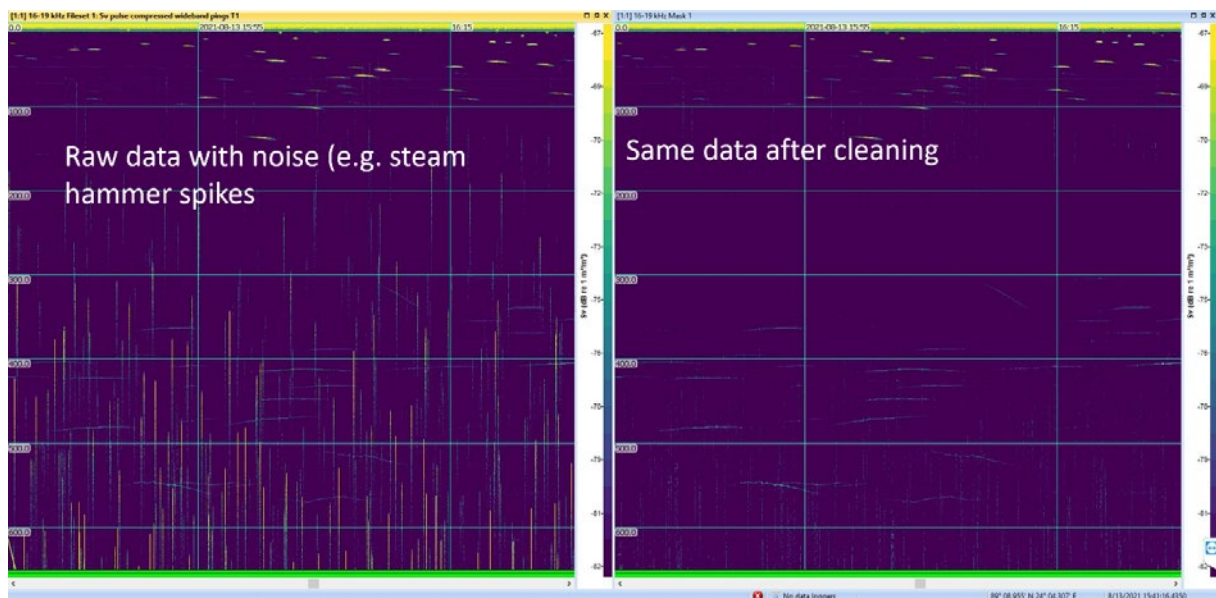
Figure 3: Cleaning of EK80 data through elimination of the steam hammer noise on IB Oden with the "Transient Noise Removal" tool Echoview™ and finalise the processing with manual data editing. The X-axis shows the depth from the surface down to 650 m, the Y-axis corresponds to time.

SAS-Oden: estimation of abundance and biomass

After cleaning of the data set was performed, the data processing followed the same procedure as for the MOSAiC data. Because the smaller targets such as copepods, amphipods and pelagic shrimps, have a size range smaller than the wavelength of 18 kHz (i.e., smaller than 8 cm), their acoustic returns are substantially weaker than larger (fish-sized) targets. Therefore, a -65 dB threshold for the single targets (means from the target tracks) and -70 dB for the echo integration was used to exclude any significant non-fish organism.

SAS-Oden: processing of the WBAT data

The WBAT equipment was attached to the CTD rosette, running two split beam transducers with the centre frequencies of 38 kHz (ES38-18DK) and 333 kHz (ES333-7CD). The 38 kHz data from the WBAT contains information on the vertical distribution of the larger organisms such as fish and macro-zooplankton (e.g., pelagic shrimps, large amphipods, gelatinous organisms), while the 333 kHz data focuses on smaller organisms (see WP3). The 38 Khz transducer was configured with an observation range of 200 m from the transducer. The 333 kHz transducer was mounted facing sideward and was configured with an observation range of 50 m. All WBAT profiles collected during this expedition were processed.

WBAT data processing required a pre-processing step for both frequencies because the configuration in the two frequencies were different (e.g., the 38 kHz was facing downward and the 333 kHz sideward). For this, the data processing platforms "R" and Matlab™ were used while the main acoustic data processing was carried out with Echoview™. The data processing included aligning of the acoustic recordings with the CTD recordings based on the ping time, generating echograms per each readjusted WBAT cast, removing the interference and background noise, removing the unwanted sections such as surface reflections, false bottom echoes, detecting the individual targets by target-tracking algorithm, and exporting the echo-integration (**Figure 4**).

Only the 38 kHz transducer was deployed with the purpose of observing fish, and in this section data processing will only be described for this transducer. An Echoview™ file was generated for each separate CTD cast was generated, pulse compressed echograms were created for the bandwidth 35-44 kHz, and the calibration parameter for this bandwidth was applied.

Depth assignment to the echograms: While the data collected with a fixed 200-m downward range from the transducer, actual depth continuously changed as the CTD was descending and ascending. Although the WBAT did not have a pressure sensor to account for depth change, the synchronised measurements were taken from the CTD based on time with a precision allowing for 0.01 m resolution. In order to import the depth information into the WBAT data analysis, a CSV line file was created in Echoview™ format and imported as a heave line into Echoview to transfer the depth information to the echograms. This resulted in V shaped echograms. For the analyses only the down-casts were used because the upcasts could be slightly biased due to the disturbance of the water column could be problematic both for echo integration and echo counting (**Figure 5**).

Data cleaning: Similar to the other acoustic data sets, a suite of noise-cleaning operations with adjusted parameters were used together with manual processing.

Target tracking: Although the ship was considered stationary during the deployments some slow drift with ice with speed below 1 knot typically occurred. To account for this, the GPS data from the ship's EK80 was used as input for the WBAT echograms. Similar to the CTD clock, the ship's EK80 clock was also synchronised to the WBAT clock. With the depth and GPS correction, target tracking was possible in the same way as with the hull mounted echosounders.

Gridding: Echo-integration results were exported in two different formats: (a) 1 minute, 10-m depth intervals, and (b) full horizontal extent, 10-m depth intervals. Only data from the downcast sections were used.
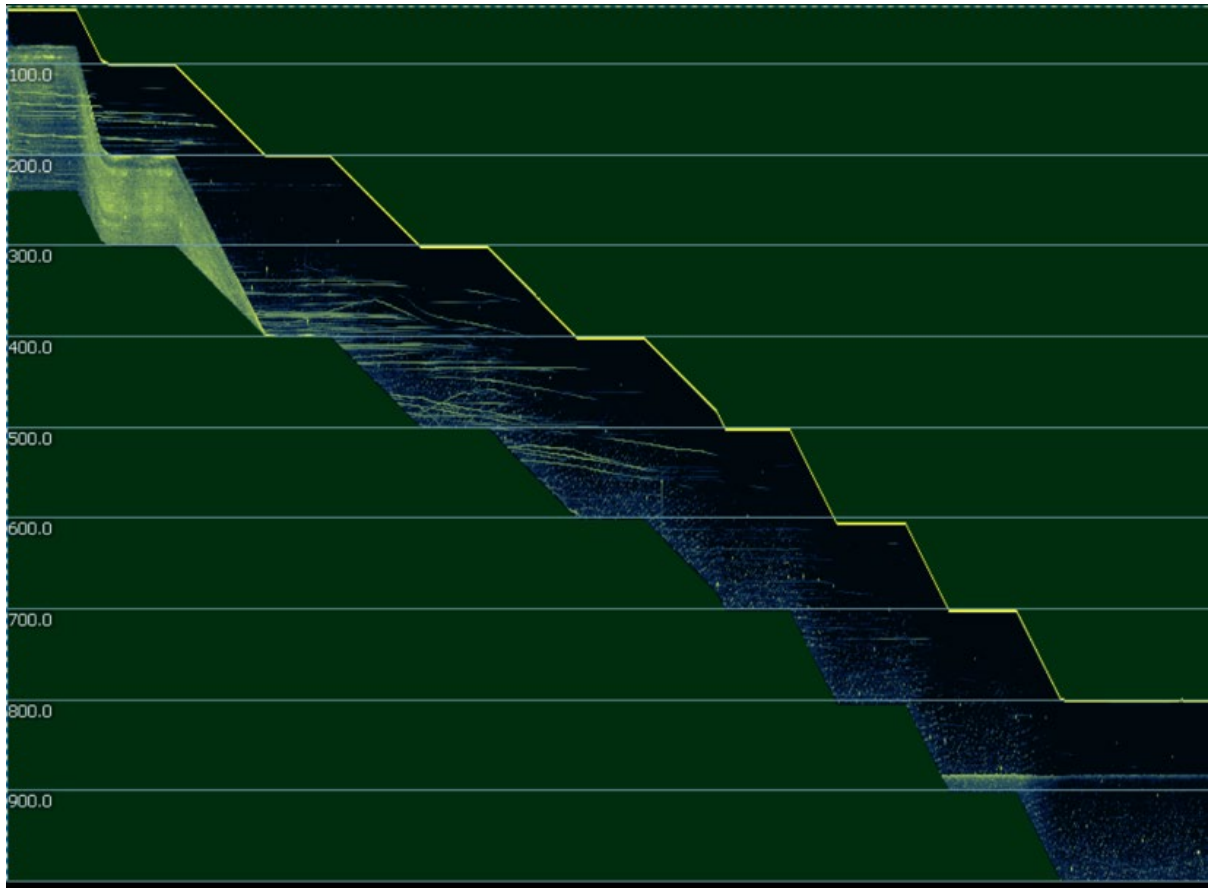
Figure 4: Example WBAT profile from one station after assigning depth to each ping. The zigzag pattern is due to regular stops every 100 m during the down cast in order enhance the target detections. The weak horizontal lines between 300 and 600 m originate from individual targets mostly very small fish.
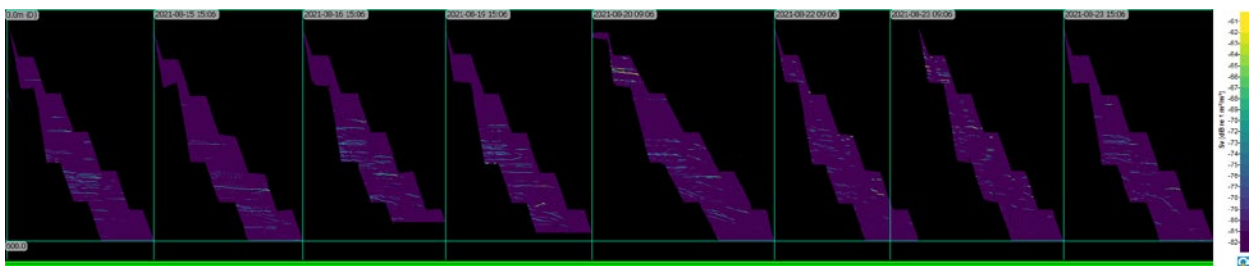
Figure 5: Example of WBAT profiles from different stations from 2021-08-15 to 2021-08-25. Each diagonal shape is representation of a separate deployment where recordings extend from surface down to 1000 m. These are cleaned echograms where eliminated data are in dark. The faint blue/yellow marks in the centre of the diagonal shapes are the detected backscatter from the mesopelagic layer between ca. 300 and 600 m of depth.

2.5. Results and discussion of WP2

MOSAiC: results from EK80 hydroacoustic data screening

After eliminating unusable pings/sections due to noise or other factors, a set of valid samples was obtained for further analysis of the MOSAiC data. Each sample represents one integration point of a 20-meter vertical and 1-hour horizontal grid. If there was no loss of data after noise cleaning, there were 720 samples per day (30 depth intervals × 24 hours). Each sample consists of a large number of data points, on average 20500 pings × 4020 vertical bins. The results indicate that high-quality acoustic data were almost continuously available throughout the entire MOSAiC year, with the exception of a two-week gap during the transit between Leg 3 and Leg 4 due to the COVID-19 pandemic (**Figure 6**).

Figure 7 shows a coloured matrix of valid samples used for the analyses, where colours indicate the relative intensity of the acoustic backscatter for each sample, altogether consisting of 344 days with valid samples from 1 October 2019 until 1 October 2020. The analyses were split into groups: Legs 1-5, steaming during transit between Legs 4 and 5, and steaming during transit home after Leg 5 until reaching the marginal ice zone. The number of valid samples during the expedition legs contain: Leg 1, 72 days, Leg 2, 71 days, Leg 3, 86 days, Leg 4, 60 days, and Leg 5, 30 days. The transit between Legs 4 and 5 contains valid samples for 11 days, and the transit home after Leg 5 contains valid samples for 12 days. **Figure 7** shows the valid data points per hour. It can be seen that the MOSAiC drift track contains almost uninterrupted data while there are large gaps during the transit and return stages of the ship. Therefore, the main analysis was focused on the track with continuous drift. **Figure 8** shows the valid acoustic measurements along the MOSAiC expedition route.
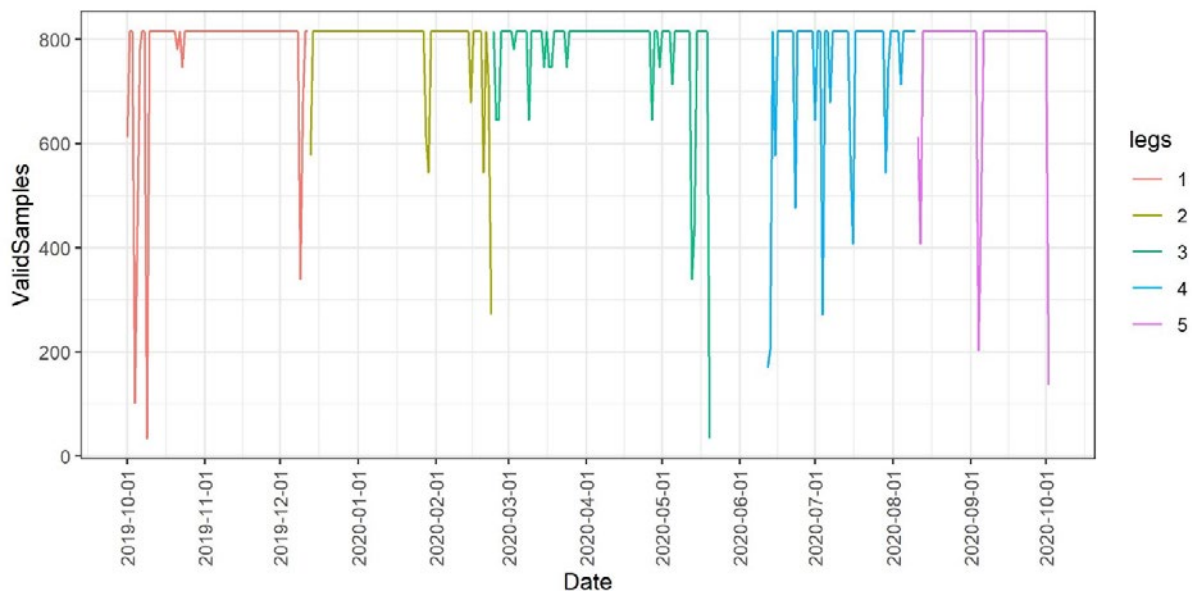
Figure 6: Valid samples per day for the 38 kHz transducer data after the noise cleaning for the MOSAiC drift expedition. Each data point represent 24 hours and number of samples are the valid grid size of 1 hour horizontal and 20 m vertical. The large gap around 1 June is due to transition between MOSAiC Legs 3 and 4 while the vessel was breaking ice.
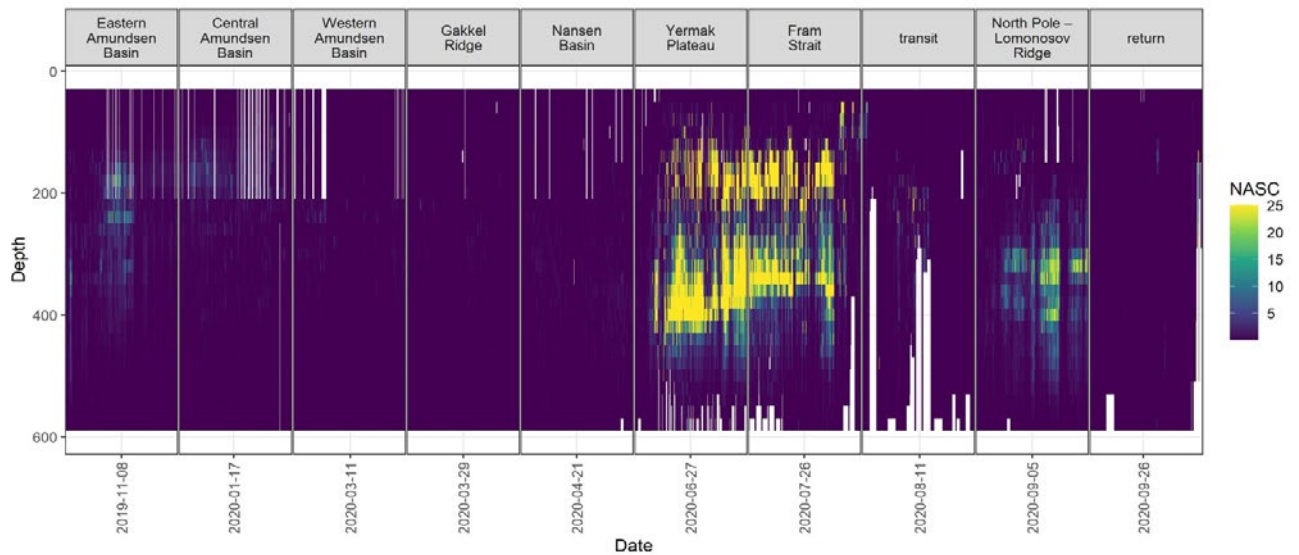
Figure 7: Final filtered/cleaned volume backscatter measured along the MOSAiC expedition route. Each pixel on this figure corresponds to a sample that is used in the further analysis. Each of these samples consist of average of data points per hour x 20 m window. White sections show the regions where no data is available due to noise. Each panel shows the geographical section of the data. The colours indicate the NASC m² km⁻², the acoustic backscatter density in that sample. For the colours, an upper cap of 25 dB was applied in order to visualise the gradients while avoiding very high density regions saturating the colour palette. The underlying range is characterised by median = 0.02 m² km⁻², mean = 10 m² km⁻², max = 24 000 m² km⁻².
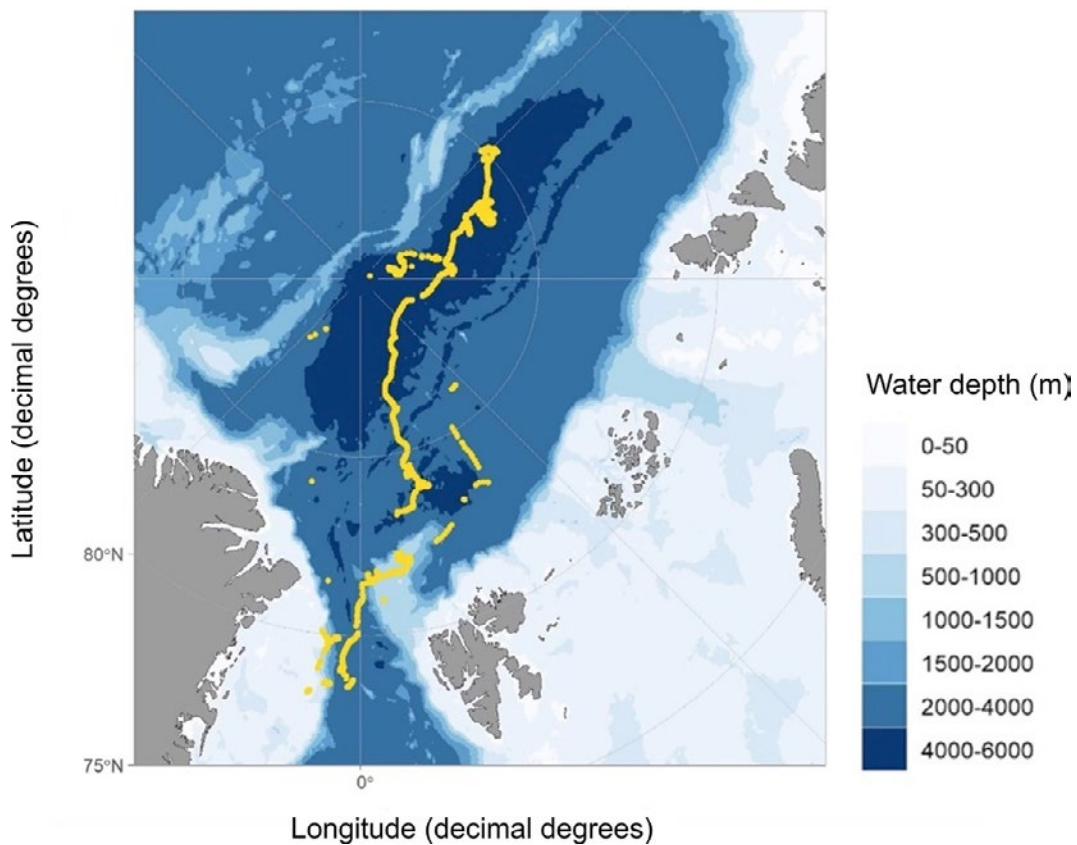
Figure 8: Valid acoustic measurements along the MOSAiC expedition route. Yellow dots show the data points where the data quality passed the noise test.

33

SAS-Oden: results from EK80 hydroacoustic data screening

Despite that large sections were omitted as a result of noise cleaning the quality of the remaining data was sufficiently high in 36 stations when the IB Oden was stationary for a longer time. Further good quality data were collected in locations when the icebreaker was stationary for other reasons (e.g., ad-hoc stops especially for acoustic recordings) and open water sections between the ice floes where the noise level was low. Altogether in 38 different locations along the expedition route, representative data was collected down to 800 m (**Figure 9**). On average 10 000 good pings were collected per location, which corresponds to more than 12 hours of measurements. Altogether, 380 817 valid pings were available for the analysis for the entire expedition route after noise cleaning.
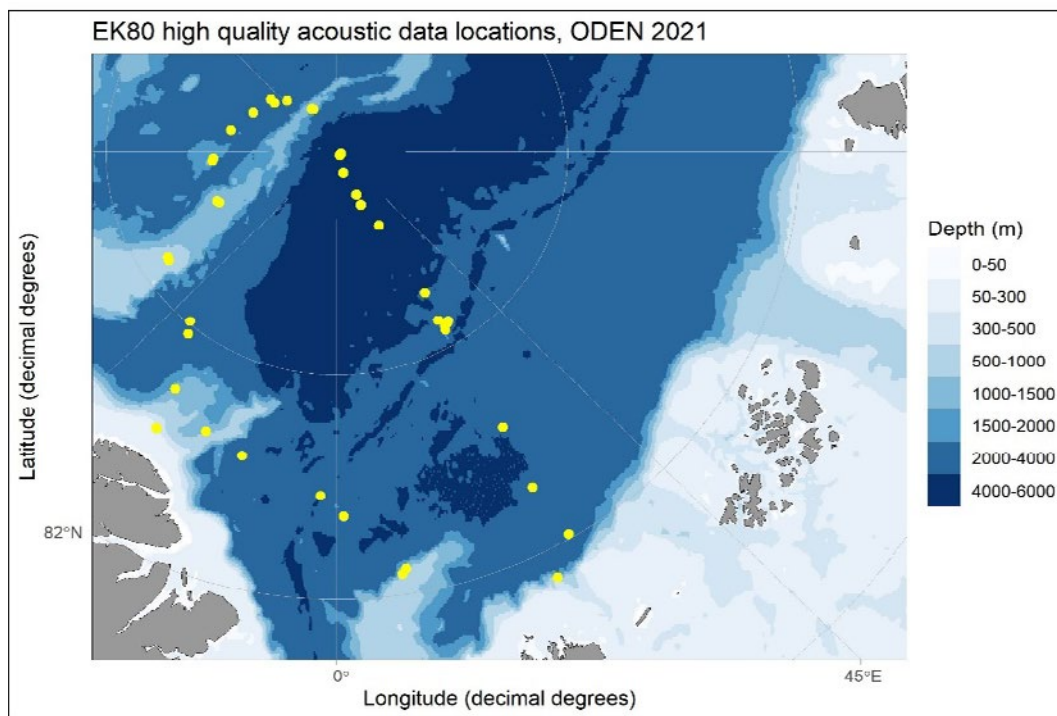
Figure 9: Positions of the 38 locations where good-quality 18 kHz EK80 data were collected down to 800 m when the icebreaker was stationary during the SAS-Oden expedition in 2021.

SAS-Oden: results from WBAT hydroacoustic data screening

Among the 31 usable 38 kHz WBAT profiles, 30 profiles covered the water column down to 600 m and 27 of them extended down to 800 m and below. Eleven of the profiles were affected by the secondary bottom reflections (aliasing) at varying levels, but only in five stations this effect interfered with the DSL at 200-600 m of depth. The remaining 26 profiles provided representative recordings for the mesopelagic zone with very high quality and resolution (**Figure 10**). The data quality near the surface was more problematic. Due to back reflections from the surface, very limited data from the upper 0-100 m depth range passed the noise cleaning, making this section not usable for target tracking.
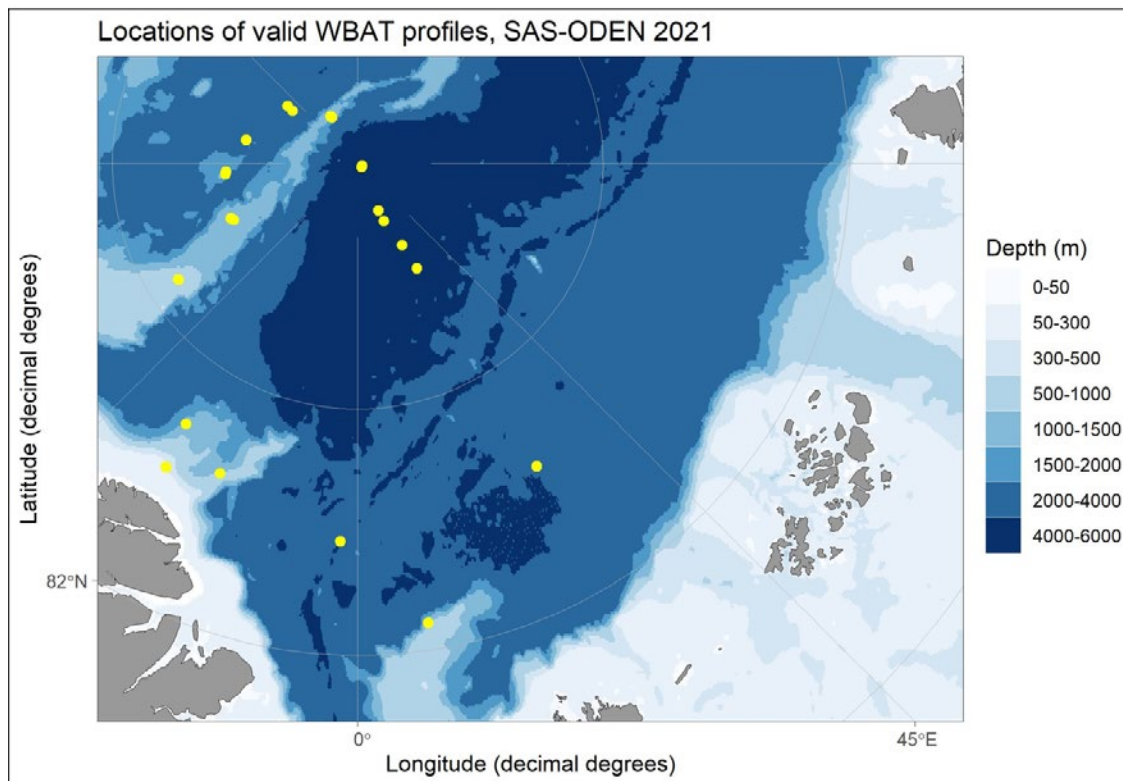
Figure 10: Positions of the 26 locations where good-quality 38 kHz WBAT data were collected down to 600 m during the SAS-Oden expedition in 2021.

MOSAiC: fish distributions from hydroacoustic data

Throughout the MOSAiC drift route, acoustic tracks of individual targets indicative of various size groups of fish were detected and changing densities of acoustic backscatter were recorded mostly as weak reflections from the mesopelagic zone (200-500 m). After carefully eliminating the non-fish targets based on the frequency responses and target-tracking analysis (see methods section), backscatter from a fish mixture of different species and size groups remained. A reliable abundance estimate from such bulk measurement can be made by using representative fish samples collected by trawling from the observed water layers. Without such information (as is the case for the SC07 project), an abundance estimation can be made by analysing individual targets to obtain an approximate indication of fish size distributions. The abundance estimates are based on the target-tracking analyses providing size distributions. The Nautical Area Scattering Coefficient (NASC) was then converted to quantities providing abundance and biomass estimates for potential species/size groups based on assumptions that are as realistic as possible. Furthermore, the number of detected individual targets provided a cross-check for the estimated numbers.

Figure 11 shows the density distribution of the bulk occurrence of fish-like targets along the expedition route of MOSAiC. The averages of the different sections are provided in **Table 12** for the conservative Target-tracking Data Set and the comprehensive Full Data Set, respectively. The highest NASC values were observed during the Leg 4 when the average NASC was 1182 m^2 km^{-2} in the Target-tracking Data Set and 1763 m^2 km^{-2} in the Full Data Set, with the latter being substantially higher than the previous. This is due to densely packed fish aggregations observed during Leg 4, which disabled reliable target detections. The lowest NASC values were observed during Leg 3 where the MOSAiC drift

was in the Amundsen and Nansen Basins. Towards the Yermak Plateau the backscatter increased dramatically and Leg 3 was separated into two parts: Part 1 with very low backscatter typical of the deep basins and Part 2 with higher backscatter when approaching the continental shelf. The NASC values were higher in both Legs 1 and 5, close to the North Pole and in the eastern part of the Amundsen basin. During Leg 2 the NASC gradually decreased as the drift proceeded eastward and in February, reaching the lowest levels just before the exchange of scientists and crew between Legs 2 and 3 around latitude 88 °N.
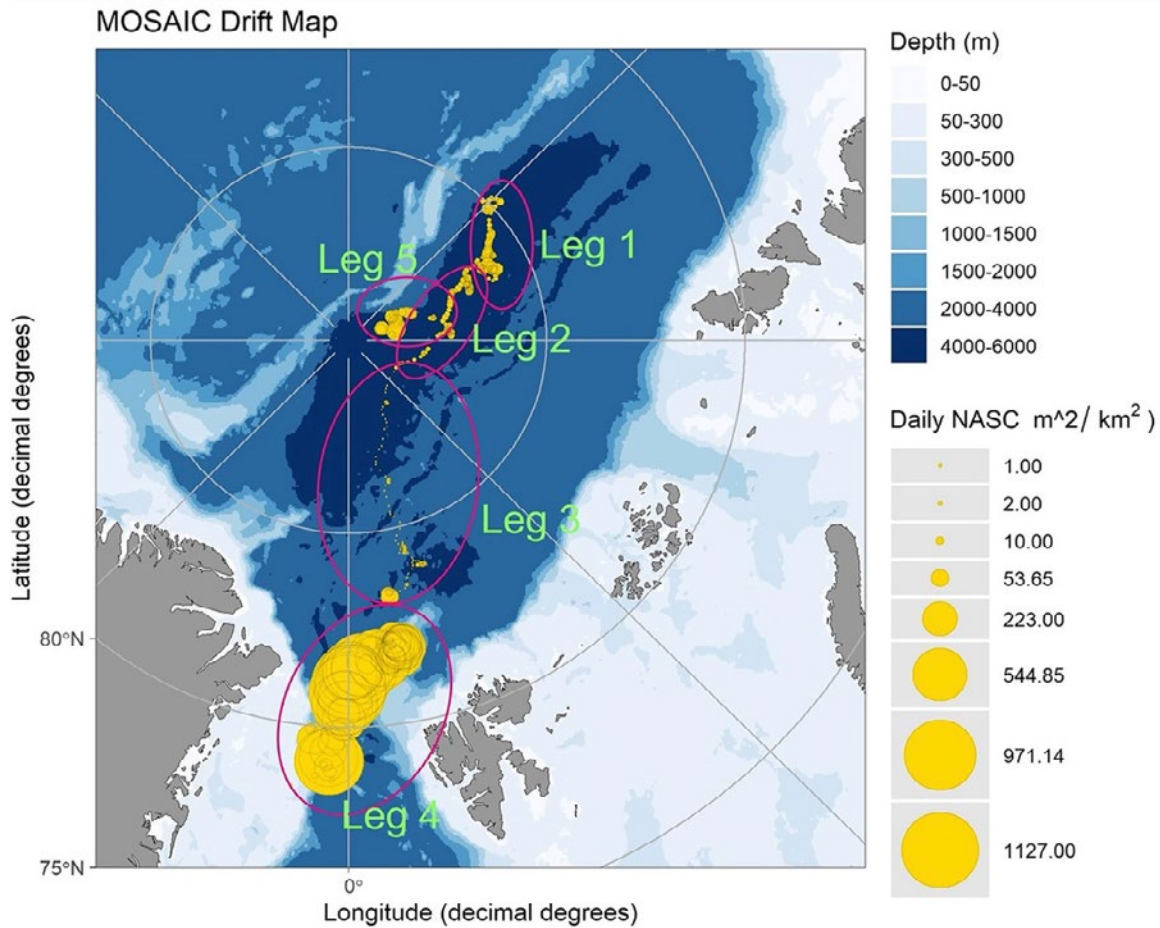
Figure 11: Nautical Area Scattering Coefficient (NASC) distribution, a metric derived from the volume backscattering density, showing the bulk occurrence of fish-like targets along the expedition route of MOSAiC.

Table 12: Average Nautical Acoustic Scattering Coefficient (NASC) in m² km⁻² for different sections of the MOSAiC expedition for the 20-600 m depth interval.

MOSAiC leg	Sub-regions	NASC Target-tracking Data Set	NASC Full Data Set
Leg 1	Eastern Amundsen Basin	16.3	19.5
Leg 2	Central Amundsen Basin	9.6	10.9
Leg 3, part 1	Western Amundsen Basin, Gakkel Ridge, Nansen Basin	1.0	1.3
Leg 3, part 2	Yermak Plateau	17.5	378.9
Leg 4	Yermak Plateau/ Fram Strait	1182.7	1763.6
Leg 5	North Pole – Lomonosov Ridge	56.9	72.3

MOSAiC: Target strength (TS) from single targets

Throughout the MOSAiC expedition route, individual targets were detected with high resolution allowing for splitting the targets into different size groups. From the target tracks, it was possible to calculate the number of individual targets for most of the MOSAiC drift route and to estimate the size of the targets. **Figure 12** shows the overall distribution of the mean TS along the MOSAiC drift track. The highest TS values were observed during Leg 4, where the drift path travels around Yermak plateau and Fram strait. In this shelf area the largest longline catches during MOSAiC were made, consisting of mostly large gadoid species such as haddock and Atlantic cod as well as beaked red fish. During Leg 4, the mean TS rose above -40 dB. The highest values reaching just above -30 dB, suggesting the presence of large gadoids up to ca. 80 cm in length. Similar observations were made shortly before the end of Leg 3.

The second-highest TS values during MOSAiC were observed in the North Pole / Lomonosov Ridge area during Leg 5. Here, the average TS was around -43 dB indicating the presence of large fish such as gadoids, but also smaller organisms, potentially myctophids and large crustaceans. There seems to be similarity in the mean TS and echogram patterns between Leg 5 and Leg 1. Both show a peak around -40 dB suggesting the presence of relatively few large gadoids, as well as a peak around -55 dB most likely due to presence of mesopelagic myctophids such as *Benthosema glaciale*, while the majority of the targets in-between could potentially be polar/Arctic cod. As for the TS signature of larger gadoids, some indication was observed during Leg 1 (October-November 2019) on the Siberian side of the Amundsen Basin. During Leg 2, the TS dropped below -50 as RV Polarstern drifted towards Greenland in the Amundsen Basin, and the patterns on the echograms indicate that small mesopelagic fish and other weakly scattering organisms start to dominate the water column. During Leg 3, as the drift further proceeds towards the Gakkel Ridge and Nansen basin, almost no large TS is observed and the mean peaks around -55 dB where myctophid type mesopelagic species are possibly becoming dominant.
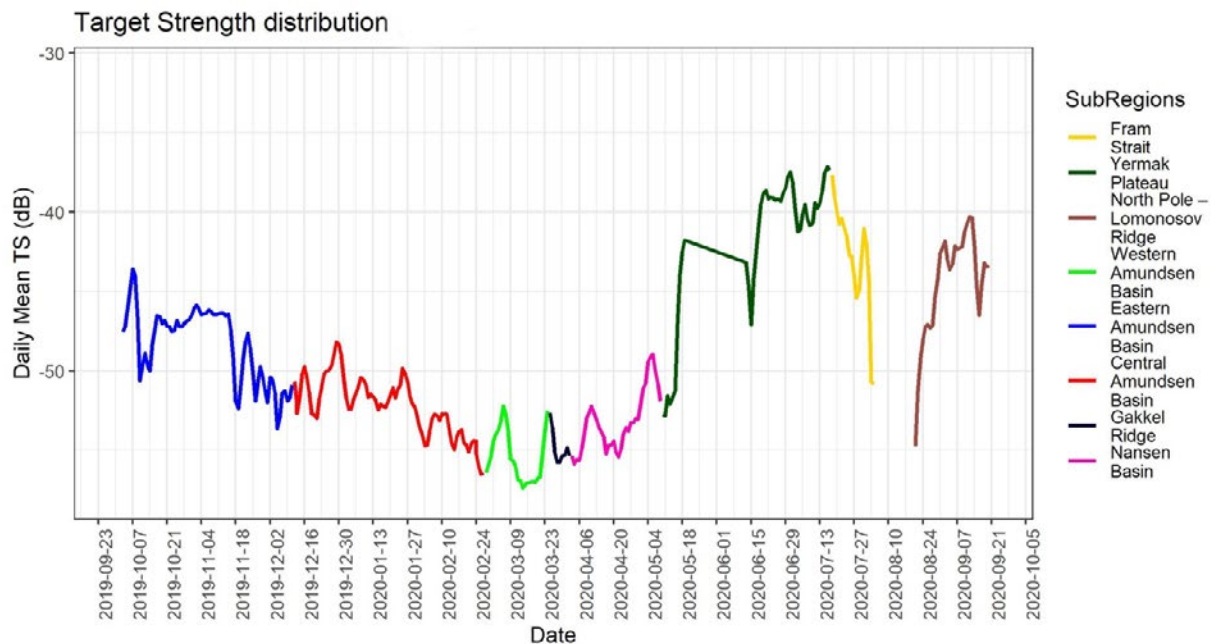
Figure 12: Distribution of the target strength (TS) for the 100-600 m depth interval. The mean TS shown here are from the fish track detections averaged per day.

MOSAiC: hypothetical abundance estimation based on acoustic measurements

From an acoustic survey with IB Oden in 2016 an initial estimate for fish abundance in the CAO was provided assuming that the Arctic endemic gadoids, polar cod *Boreogadus saida* and ice cod *Arctogadus glacialis* are the most realistic candidate species based on historical recordings from the region and the known physiology of these species[44]. This was based on the previously mentioned assumptions for these species, while also testing for the larger North-Atlantic species that were captured during the MOSAiC drift: Atlantic cod *Gadus morhua*, haddock *Melanogrammus aeglefinus*, and beaked redfish *Sebastes mentella*.

In addition to these relatively larger fish, there is evidence for the presence of the small myctophid *Benthosema glaciale* in the CAO both from MOSAiC (see WP3) and from a recent Norwegian expedition with RV Kronprins Haakon in 2021[45]. During the latter expedition both hydroacoustic and trawl sampling took place in late summer. Their 12 trawl hauls, partially overlapped in space both with the MOSAiC drift 2019-2020 and the SAS-Oden expedition. They captured seven individuals of three different species: two polar cod *Boreogadus saida*, four glacier lanternfish *Benthosema glaciale* and one (small juvenile) Greenland halibut *Reinhardtius hippoglossoides.* The presence of the small myctophid *Benthosema glaciale* justifies the chosen -65 dB lower threshold for the target strength (**Figure 13**) and constitutes the best possible candidate for the lowest-density acoustic backscatter regions of MOSAiC.

For Legs 1, 4, and 5 of MOSAiC a secondary peak between 20 and 40 cm is also distinguishable (**Figure 13**). This group could be an indication of adult polar cod (ca. 25 cm in length), while containing some North Atlantic gadoids (larger than ca. 30 cm) as well. The dominance of the adult cod at this layer is in line with the hypothesis suggesting the possible descent of *Boreogadus* to the mesopelagic zone of the CAO when they grow older to exploit the zooplankton there. The part of the histogram in Leg 4 where the size distribution goes beyond 40 cm, it can be assumed that the large North Atlantic gadoids and beaked red fish dominate the distribution.

MOSAiC Scenario 1: Boreogadus saida and Arctogadus glacialis dominance

Scenario 1 assumed that *Boreogadus saida*[46] and ice cod *Arctogadus glacialis*[47] were the only two possible candidates for the entire region using the MOSAiC data. A published TS/length relationship[48] was used followed by conversion of the TS measurements into length distributions: TS = 21.8 log10 L – 72.7[49]. The weight of the hypothetical fish individuals was calculated using the equation $0.0053*(10^{((TS+72.7)/21.8)})^{3.042}$, where 0.0053 and 3.042 represent a and b in the length/weight relationship $(a*L^b)$[50]. The estimated average lengths for this scenario are presented in **Figure 14**. According to these results, average sizes as small as 6.7 cm (Leg 3) and as large as 22.1 cm would occur. **Figure 15** gives a more detailed distribution from the detected tracks. In all

[44] Snoeijs-Leijonmalm P, et al. (2021) A deep scattering layer under the North Pole pack ice. Progress in Oceanography, 194:102560 [https://doi.org/10.1016/j.pocean.2021.102560]

[45] Ingvaldsen RB, et al. (2023) Under-ice observations by trawls and multi-frequency acoustics in the Central Arctic Ocean reveals abundance and composition of pelagic fauna. Scientific Reports 13:1000 [https://doi.org/10.1038/s41598-023-27957-x]

[46] Mueter FJ, et al. (2016). The ecology of gadid fishes in the circumpolar Arctic with a special emphasis on the polar cod (*Boreogadus saida*). Polar Biology 39:961-967 [https://link.springer.com/article/10.1007/s00300-016-1965-3]

[47] Jordan AD, et al. (2003). Revision of the Arctic cod genus *Arctogadus*. Journal of Fish Biology 62:1339-1352 [https://doi.org/10.1046/j.1095-8649.2003.00115.x]

[48] Snoeijs-Leijonmalm P, et al. (2021) A deep scattering layer under the North Pole pack ice. Progress in Oceanography, 194:102560 [https://doi.org/10.1016/j.pocean.2021.102560]

[49] Mamylov VS (2003) About the comparison of fish distribution densities estimated using trawl and acoustic methods. Report from the Improvement of instrumental methods for stock assessment of marine organisms, Russian-Norwegian Workshop, 11–14 November 2003, PINRO, Murmansk, Russia

[50] Fey DP & Węsławski JM(2017). Age, growth rate, and otolith growth of polar cod (*Boreogadus saida*) in two fjords of Svalbard, Kongsfjorden and Rijpfjorden. Oceanologia 59:576-584 [https://doi.org/10.1016/j.oceano.2017.03.011]

sections a peak around 5-10 cm is visible. Only in Leg 3 this peak is a bit more shifted towards 5 cm. The fact that this peak indicates smaller sizes than the other legs could be an indication of possible dominance of another small-sized taxon, e.g., myctophids such as glacier lanternfish.

Figure 14 suggests that the average size distribution from the second week of October 2019 until May 2020 was below 20 cm, which confirms the expected size distribution of polar cod if the spatial distribution of the adult population differs from the juveniles. In this Scenario 1, adults would mainly be concentrated in the eastern Amundsen basin and the North Pole area whereas juveniles are spread in the entire CAO. **Figure 15** shows the potential average weight according to this scenario. While the average weight would be around 400 g per individual in the eastern Amundsen basin during Leg 1, towards the west, it drops down to 100 g, and towards the Gakkel ridge and Nansen basin it goes below 50 g. The reason that the average weight during Leg 3 is larger than that in Leg 2 would again be the changed composition towards the end of Leg 3 near Yermak plateau. If this section is excluded, the mean weight goes down to 22 g for the first part of Leg 3. Data collected during Leg 5, while located at similar high latitudes as Leg 1, shows larger numbers in the estimated length and weight distributions, where the average weight per individual goes above 500 g. One explanation this could be dominance of *Arctogadus* rather than *Boreogadus* because the former species is known to grow older and larger. The biomass estimates resulting from the calculations are shown in **Figure 16** and **Table 13**.
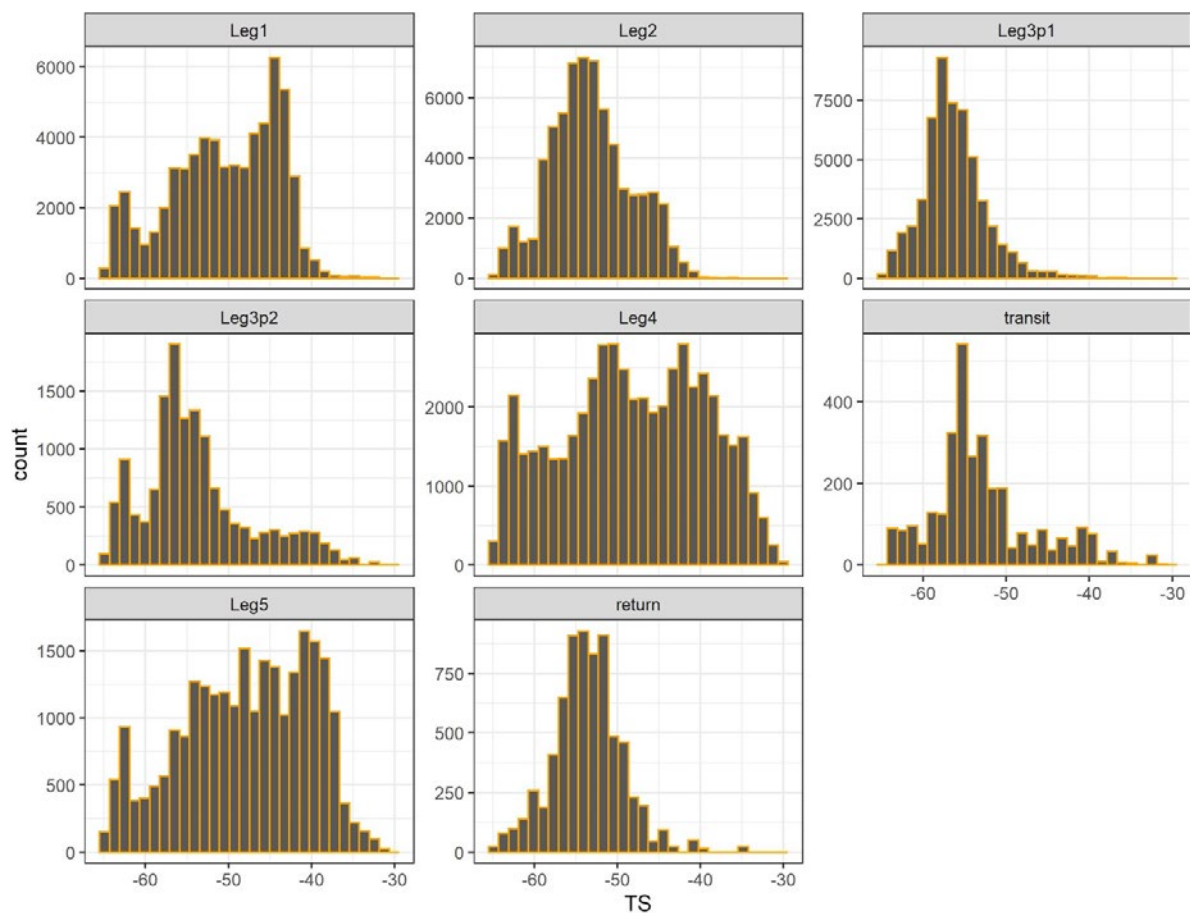
Figure 13: Distribution of the target strength (TS) for each MOSAiC section. The multiple peaks in Legs 1, 4, and 5 are potentially due to a mixture of different taxonomic groups such as gadoids and small mesopelagic fish such as myctophids.
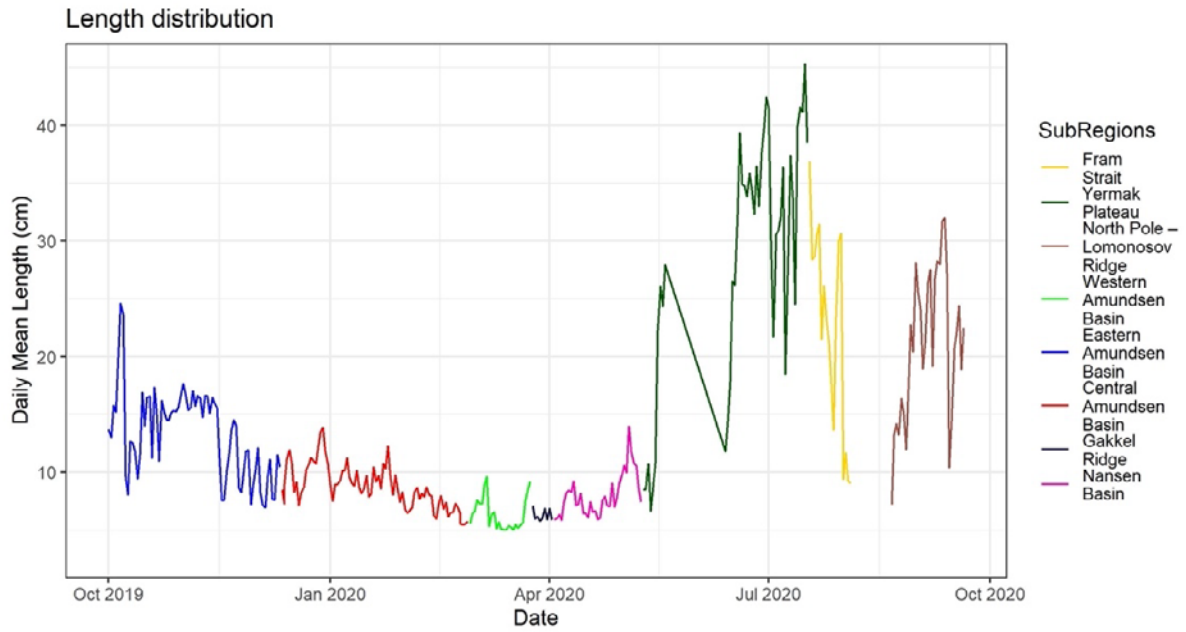
Figure 14: MOSAiC Scenario 1: stimulated length distribution calculated from the fish-track-based target strengths assuming *Boreogadus* and *Arctogadus* dominance. Data points indicate the averages for every 6 hours of the fish tracks detected below 200 m.
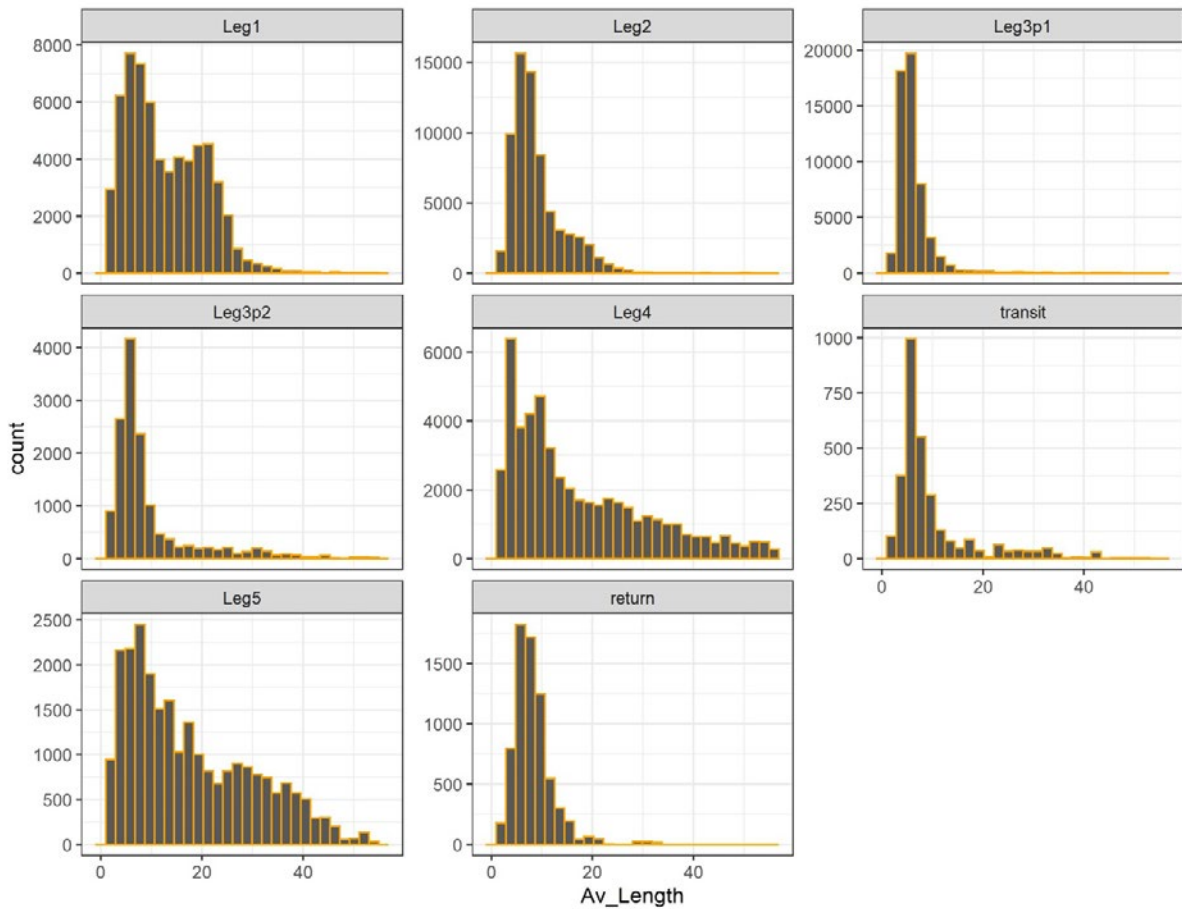
Figure 15: MOSAiC Scenario 1: estimated Length distribution calculated from the fish track based target strengths assuming *Boreogadus* and *Arctogadus* dominance.
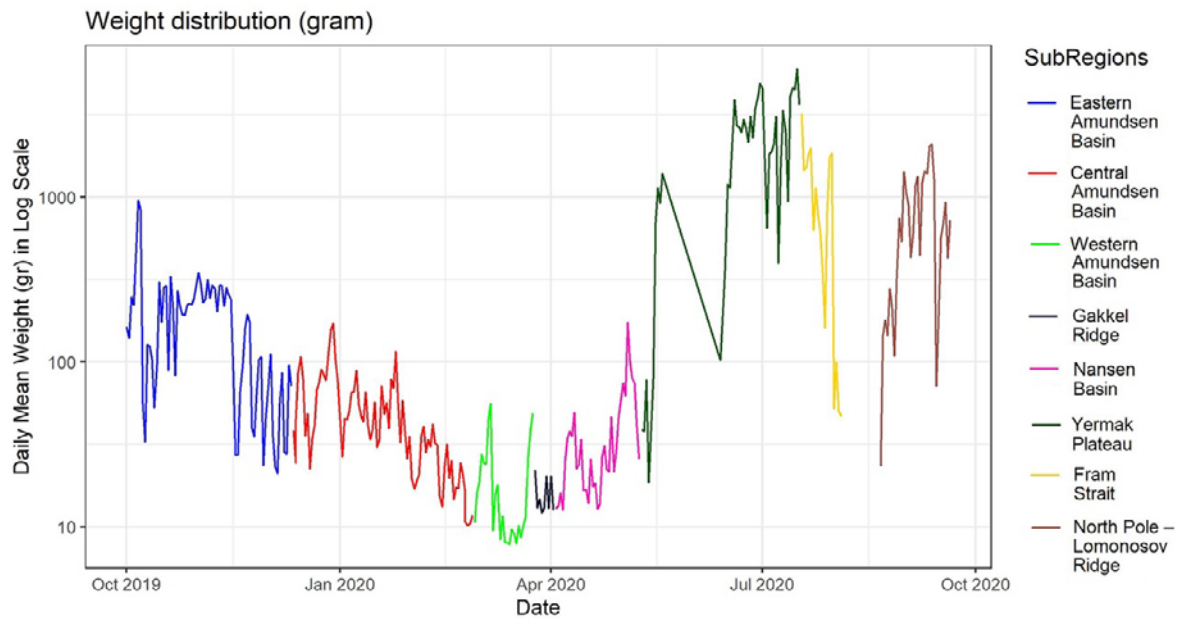
Figure 16: MOSAiC Scenario 1: Estimated average weight for given length distribution for each 6-hour interval for individual fish assuming *Boreogadus* and *Arctogadus* dominance. The Y-axis is on a Log10 scale and weight is given in gram.

Table 13: MOSAiC Scenario 1: results of estimates for mean biomass, abundance, fish length and fish weight assuming *Boreogadus* and *Arctogadus* dominance for different sections in the deep scattering layer (DSL) along the MOSAiC route. The mean area scattering coefficient (NASC) and mean target strength (TS) include measurements from the 20-600 m depth interval during the MOSAiC expedition. Results are given for both for the Target-tracking Data Set and the Full Data Set providing lower and upper boundaries of the estimates, respectively.

Data set used	MOSAiC leg	Mean NASC (m² km⁻²)	Mean TS (dB)	Mean biomass (kg km⁻²)	Mean number (ind. km⁻²)	Mean length (cm)	Mean weight (g)
Target-tracking Data Set	Leg 1	4.8	-48.2	121	11705	10.9	21.4
	Leg 2	2.8	-51.8	48	12883	7.9	6.3
	Leg 3, part 1	0.3	-54.2	6	2058	6.0	3.5
	Leg 3, part 2	7.1	-46.6	352	4877	11.3	43.2
	Leg 4	328.2	-41.1	21333	222095	21.3	217.4
	Leg 5	15.6	-43.1	718	9642	18.8	102.3
Full Data Set	Leg 1	5.7	-47.3	145	14666	12.0	28.2
	Leg 2	3.2	-51.2	54	14793	8.4	7.6
	Leg 3, part 1	0.4	-54.0	7	2883	6.1	3.7
	Leg 3, part 2	8.4	-46.4	376	8903	11.2	47.4
	Leg 4	472.2	-41.2	23937	1173064	20.9	208.5
	Leg 5	19.0	-43.1	736	24908	18.3	107.9
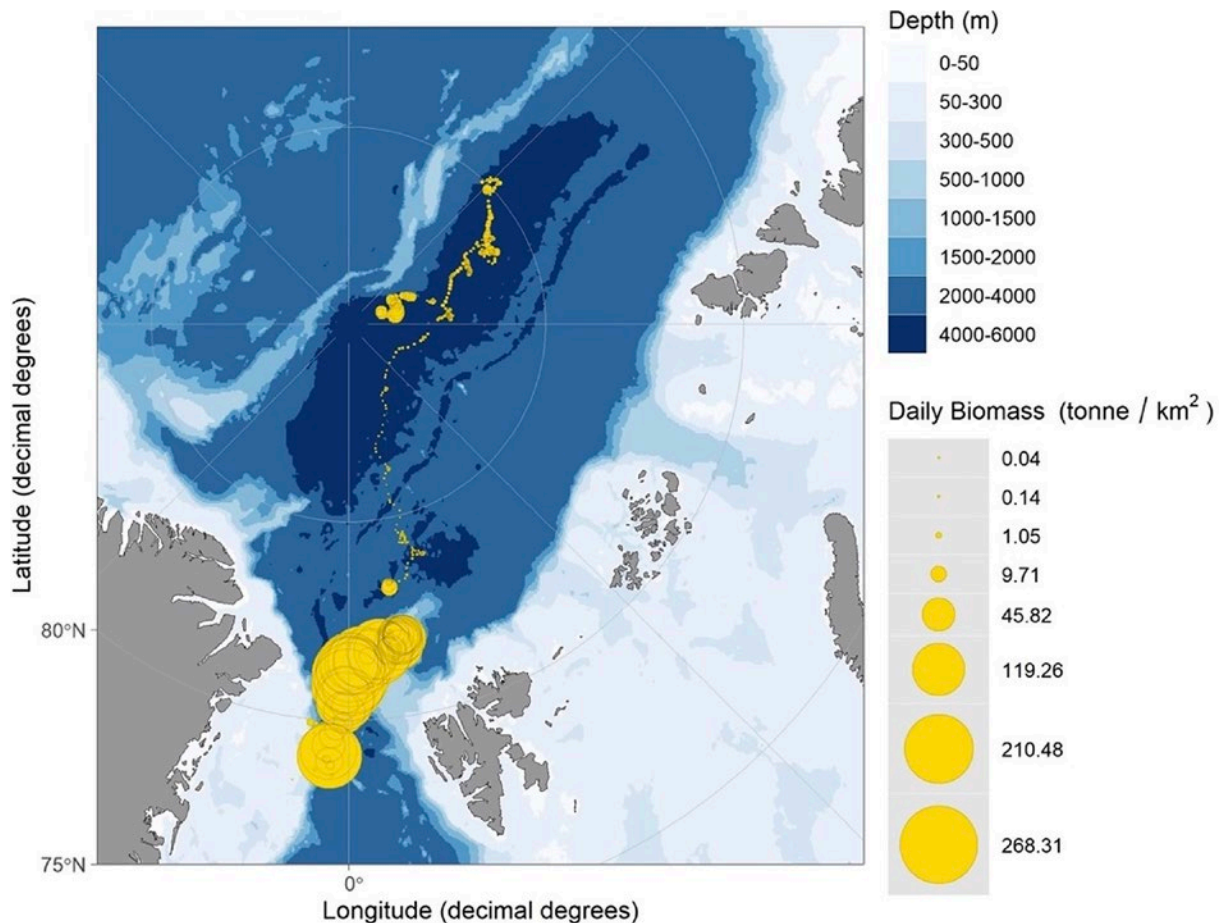
Figure 17: MOSAiC Scenario 1: biomass distribution along the MOSAiC route assuming *Boreogadus* and *Arctogadus* dominance.

MOSAiC Scenario 2: Boreogadus saida, Benthosema glaciale, squid, Atlantic gadoids

In Scenario 2 a mixture of species was assumed for the entire region using the MOSAiC data. Simulations were used to create a distribution similar to the TS values as exported from the fish tracks from MOSAiC. In this scenario it was assumed that all *Boreogadus* in the water layer below 200 m would be adults. This scenario is based on several previous studies at lower latitudes. One study in the south-eastern Beaufort Sea found *Boreogadus* total length ranging from 11.6 to 25.9 cm (mean 16.4 cm, standard deviation 2.2 cm)[51]. A similar distribution was found in the Rijpfjorden in northern Svalbard, with single females reaching up to 30.5 cm and pointing out that the fish sizes were larger in the Arctic compared to the sub-Arctic stations of the study[52]. It has been reported that *Boreogadus* can reach up to maximum 40 cm and weigh 430 g[53,54], and in Scenario 2 therefore assumed the lower length of *Boreogadus* to be 11 cm and maximum length 40 cm, with an expected mean length of 15 cm. Using the 21.8 ´ Log(L)-72.7 equation, the TS range for *Boreogadus* would be between -50 dB and -37.8 dB and the peak should be expected around -47 dB.

[51] Benoit D, et al. (2008) Hydroacoustic detection of large winter aggregations of Arctic cod (*Boreogadus saida*) at depth in ice-covered Franklin Bay (Beaufort Sea). Journal of Geophysical Research: Oceans 113:C6 [https://doi.org/10.1029/2007JC004276]

[52] Nahrgang J, et al. (2014) Gender specific reproductive strategies of an Arctic key species (*Boreogadus saida*) and implications of climate change. PLoS one 9:e98452 [https://doi.org/10.1371/journal.pone.0098452]

[53] Boitsov VD, et al. (2013) Polar cod of the Barents Sea. PINRO, Murmansk

[54] Aune M, et al. (2021). Distribution and ecology of polar cod (*Boreogadus saida*) in the eastern Barents Sea: A review of historical literature. Marine environmental research, 166, 105262 [https://doi.org/10.1016/j.marenvres.2021.105262]

For TS values below -50 dB at depths below 200 m a likely candidate would be the glacier lanternfish *Benthosema glaciale* for which a length range between 4.5 and 6 cm has been reported from a recent expedition in the CAO and peak TS at 38 kHz between -55 dB and -54 dB[55]. In a Norwegian fjord focusing on a DSL around 200-390 m of depth glacier lanternfish was a dominant fish and a range of -65 dB to -50 dB was used[56]. In other Norwegian fjords similar values were found for glacier lantern fish: TS varied from -54 to -69 dB at 38 kHz for mean fish length 5.4 cm[57] and TS varied from -62 to -52 dB for mean fish length 6 cm mean length[58]. These values are confirmed by even more detailed studies taking the swim bladder size into account[59]. The latter study underlines the difficulty for modelling this fish due to variability in the gas-filled swim bladder dimension as the adults gets older. They fill their swim bladder with lipids which nonlinearly changes the resonance characteristics and the TS and it is difficult to establish a length/TS relationship without knowledge of swim bladder conditions. In this study parameterization for *Benthosema glaciale* was performed based on the histogram analysis of the TS distribution as well as on parameters used in other studies using fish length 4.5-6 cm as peak range and using 10 cm as maximum length.

During MOSAiC TS values higher than the largest possible *Boreogadus* (above -37.8 dB) were recorded, suggesting presence of larger fish such as Atlantic gadoids. The fish catches during MOSAiC also confirm this assumption (see WP5). For this, the typical gadoid equation from Foote[60] for Physoclisti species, $20*\text{Log}(L)-67.5$ was used. As for the length/weight relationship, $0.0059*L^{3.11}$ was used[61]. In general, there remained one group peaking between -50 dB and -55 dB that could not be explained neither by *Boreogadus* nor by *Benthosema.* For this, one last candidate (but tentative due to its presumed very low scattering abilities), armhook squid *Gonatus fabricii*, observed by the FishCam video recordings (see WP3), was used. For the TS equation, in-vitro measurements of Kang et al.[62] for the squid *Todarodes pacificus* ($20*\text{log}(ML)-74.5$ where ML = mantle length) and a published length/weight relationship ($0.0004*ML^{2.4705}$)[63] were used.

The results of Scenario 2 are shown in **Figure 18** and **Figure 19** as the computed distributions that mimic the TS distributions of the different sections of the MOSAiC expedition. Due to a changing community composition during different sections of the transpolar drift, the TS distributions were different at each section, and depending on the complexity some of the distributions were multimodal. Therefore, assumptions were not only made for taxonomic groups but also for mean length with standard deviation (SD) in order to match the measured TS.

[55] Ingvaldsen RB, et al. (2023) Under-ice observations by trawls and multi-frequency acoustics in the Central Arctic Ocean reveals abundance and composition of pelagic fauna. Scientific Reports 13:1000 [https://doi.org/10.1038/s41598-023-27957-x]

[56] Dypvik E, et al. (2012). Seasonal variations in vertical migration of glacier lanternfish, *Benthosema glaciale*. Marine Biology 159:1673-1683 [https://doi.org/10.1007/s00227-012-1953-2]

[57] Torgersen T & Kaartvedt S (2001) In situ swimming behaviour of individual mesopelagic fish studied by split-beam echo target tracking. ICES Journal of Marine Science 58:346-354 [https://doi.org/10.1006/jmsc.2000.1016]

[58] Kaartvedt S, et al. (2009) Use of bottom-mounted echo sounders in exploring behavior of mesopelagic fishes. Marine Ecology Progress Series 395:109-118 [https://doi.org/10.3354/meps08174]

[59] Scoulding B, et al. (2015) Target strengths of two abundant mesopelagic fish species. The Journal of the Acoustical Society of America 137:989-1000 [https://doi.org/10.1121/1.4906177]

[60] Foote KG (1987) Fish target strengths for use in echo integrator surveys. The Journal of the Acoustical Society of America 82:981 [https://doi.org/10.1121/1.395298]

[61] Morris CJ & Green JM (2002) Biological characteristics of a resident population of Atlantic cod (*Gadus morhua* L.) in southern Labrador. ICES Journal of Marine Science 59:666-678 [https://doi.org/10.1006/jmsc.2002.1228]

[62] Kang D, et al. (2005) The influence of tilt angle on the acoustic target strength of the Japanese common squid (*Todarodes pacificus*). ICES Journal of Marine Science 62:779-789 [https://doi.org/10.1016/j.icesjms.2005.02.002]

[63] Arkhipkin AI & Bjørke H (1999) Ontogenetic changes in morphometric and reproductive indices of the squid *Gonatus fabricii* (Oegopsida, Gonatidae) in the Norwegian Sea. Polar Biology 22:357-365 [https://doi.org/10.1007/s003000050429]

During Leg 1 of MOSAiC, a mixture of *Benthosema*, *Boreogadus*, a gadoid group and *Gonatus* was assumed. *Benthosema* consisted of three length modes: mean 6 cm (SD=0.9), 5.5 cm (SD=0.7) and 3 cm (SD=0.5). *Boreogadus* consisted of two length modes: mean 15 cm (SD=4) and 12 cm (SD=1). The gadoid group consisted of one length mode: mean 16.5 cm (SD=3.5) and *Gonatus* had a unimodal distribution with mean mantle length 10.5 cm (SD=2.6) (**Table 14**). The Length distribution (mean and standard deviation for each taxonomic group) is the result of a simulation trying to mimic the TS distribution from the detected fish tracks. As there is multimodality in the TS distribution, different peaks would be coming from different species/size groups. In order to disentangle different size distributions, different peaks from the histograms were manually analysed and assigned to species based on the expected TS. At this stage, the simulation algorithm tests the parameters that could best fit the full distribution when all combined. Eventually, the best fit is selected as the most likely species/size composition.

Table 14: Summary of the groups hypothetically constituting the TS measured during MOSAiC Leg 1 in Scenario 2.

Taxon	Mean length (cm)	TS (dB)	Percentage
Benthosema glaciale (myctophid)	4.3	-58.32	16 %
Boreogadus saida (gadoid)	15.8	-46.31	28 %
Atlantic gadoids	16.5	-42.97	28 %
Gonatus fabricii (squid)	10.5	-52.84	28 %

During Leg 2 of MOSAiC, a mixture of *Benthosema*, *Boreogadus* and *Gonatus* but no other gadoids was assumed because of the absence of any conspicuous group with high TS. *Benthosema* consisted of two length modes: mean 6 cm (SD=0.9) and 5.5 cm (SD=0.7). *Boreogadus* consisted of two length modes: mean 13 cm (SD=5) and 12 cm (SD=1) and the lengths below 11 cm were filtered out assuming that this is the lower size limit for the *Boreogadus* adults. *Gonatus* had a unimodal distribution with mean mantle length 9 cm (SD=3) (**Table 15**).

Table 15: Summary of the groups hypothetically constituting the TS measured during MOSAiC Leg 2 in Scenario 2.

Taxon	Mean length (cm)	TS (dB)	Percentage
Benthosema glaciale (myctophid)	5.7	-56.33	11 %
Boreogadus saida (gadoid)	15.6	-46.43	17 %
Gonatus fabricii (squid)	9.2	-53.91	72 %

MOSAiC Leg 3, Part 1 is the section of the Leg 3 before the MOSAiC drift approaches the Yermak Plateau. This section corresponds to the lowest backscattering density measurements made during the entire expedition, where the TS values from the individual fish tracks were also smallest. These measurements allowed for only one candidate, *Benthosema,* with three length modes: mean 6 cm (SD=4), 5.5 cm (SD=4) and 5 cm (SD=1). The overall mean TS for this group was -55.82 dB and the mean length was 5.8 cm.

MOSAiC Leg 3, Part 2 covered the last two weeks of Leg 3 when the MOSAiC drift approached the Yermak Plateau. This is a relatively complex section where the large gadoids start to appear as transition from the CAO towards the shelf system starts. Here, the Atlantic influence strengthens and a suddenly changed community composition created a mixture of TS values both from small-sized and large-sized individuals. For this section a mixture of *Benthosema*, a gadoid group and *Gonatus* was assumed.

Benthosema had a unimodal distribution with a mean length of 5 cm (SD=4). The gadoid group consisted of two length modes: mean 16.5 cm (SD=5.5) and 22.5 cm (SD=8). *Gonatus* had a unimodal distribution with mean mantle length 9 cm (SD=4) (**Table 16**).

Table 16: Summary of the groups hypothetically constituting the TS measured during MOSAiC Leg 3, Part 2 in Scenario 2.

Taxon	Mean length (cm)	TS (dB)	Percentage
Benthosema glaciale (myctophid)	6.6	-54.73	46 %
Atlantic gadoids	19.7	-41.06	19 %
Gonatus fabricii (squid)	9.6	-53.32	35 %

MOSAiC Leg 4 had remarkably different distribution than the first three expedition legs where dense aggregations and large individuals were detected. In line with the observed patterns on echograms and catches with the long lines. A mixture of *Benthosema* with a mean length of 5 cm (SD=4) was assumed to account for the small-sized detections in sections where the large-sized group disappeared. The gadoid group dominated this section and three length modes were assumed: mean 40 cm (SD=9.5), 20.5 cm (SD=7.5) and 12 cm (SD=7). *Gonatus* was added to mimic the full distribution in the measured TS values to mimic a unimodal distribution with mean mantle length 11.5 cm (SD=4) (**Table 17**).

Table 17: Summary of the groups hypothetically constituting the TS measured during MOSAiC Leg 4 in Scenario 2.

Taxon	Mean length (cm)	TS (dB)	Percentage
Benthosema glaciale (myctophid)	6.6	-54.73	17 %
Atlantic gadoids	23.2	-39.25	65 %
Gonatus fabricii (squid)	11.7	-51.53	17 %

In the beginning of Leg 5 of MOSAiC, RV Polarstern departed from its southern location in Fram Strait and steamed towards the North Pole. After anchoring to a new ice floe, the expedition drifted slightly west of the North Pole around the Lomonosov Ridge. The patterns observed in the backscatter during Leg 5 resembled those of Leg 1, especially regarding the multimodality of the distribution, although with slightly higher volume backscatter and distinctly higher TS values.

During Leg 5 of MOSAiC, a mixture of *Benthosema*, *Boreogadus*, a gadoid group and *Gonatus* was assumed. *Benthosema* had a unimodal distribution with mean length 5 cm (SD=4). *Boreogadus* consisted of two length modes: mean 15 cm (SD=5) and 12 cm (SD=3). The gadoid group consisted of three length modes: mean 31 cm (SD=9.5), 17.5 cm (SD=7.5) and 12 cm (SD=7). *Gonatus* had a unimodal distribution with mean mantle length 11.5 cm (SD=4) (**Table 18**).

Table 18: Summary of the groups hypothetically constituting the TS measured during MOSAiC Leg 5 in Scenario 2.

Taxon	Mean length (cm)	TS (dB)	Percentage
Benthosema glaciale (myctophid)	6.6	-54.61	21 %
Boreogadus saida (gadoid)	14.5	-47.19	7 %
Atlantic gadoids	20.5	-40.39	54 %
Gonatus fabricii (squid)	11.6	-51.57	18 %

The gadoid group consisted of three modes similar to Leg 4, a large group with mean 31 cm (SD=9.5) a medium-sized group, mean 17.5 cm (SD=7.5) and a relatively small-sized group mean 12 cm (SD=7). A final squid group was generated, with mean mantle length of 11.5 cm (SD=4). In Scenario 2 the acoustic backscatter was converted to biomass according to **Table 19** based on combination of different assemblages of taxa and size groups. The mean TS and mean individual weight was calculated based of relative contribution of each taxonomic group. The Results of MOSAiC Scenario 2 are shown in **Table 20**, **Figure 18**, and **Figure 19**.

Table 19: Average target strength and mean weight for different hypothetical communities during the different legs of the MOSAiC expedition.

MOSAiC leg	σ_{bs}/(1 m^2) backscattering cross section	Mean weight (g)	Mean TS (dB)
Leg 1	2.09E-05	14.7	-46.80
Leg 2	7.43E-06	12.6	-51.29
Leg 3, Part 1	3.26E-06	2.2	-54.87
Leg 3, Part 2	1.69E-05	15.9	-47.73
Leg 4	7.96E-05	70.5	-40.99
Leg 5	5.02E-05	40.9	-43.00

Table 20: MOSAiC Scenario 2, results of estimates for mean biomass, abundance, fish length and fish weight assuming *Boreogadus* and *Arctogadus* dominance for different sections in the deep scattering layer (DSL) along the MOSAiC route. The mean area scattering coefficient (NASC) and mean target strength (TS) include measurements from the 20-600 m depth interval during the MOSAiC expedition. Results are given for both for the Target-tracking Data Set and the Full Data Set providing lower and upper boundaries of the estimates, respectively.

Data set used	MOSAiC region	Mean NASC (m^2 km^{-2})	Mean TS (dB)	Mean biomass (kg km^{-2})	Mean number (ind. km^{-2})	Mean length (cm)	Mean weight (g)	Mean date
Target-tracking Data Set	Eastern Amundsen Basin	4.8	-48.1	123	11669	11	21.7	191106
	Central Amundsen Basin	2.7	-52.0	46	12436	7.7	6	200121
	Western Amundsen Basin	0.4	-54.5	7	2761	5.7	3	200312
	Gakkel Ridge	0.1	-55.4	2	615	5.4	2.4	200329
	Nansen Basin	0.5	-51.9	10	2263	7.2	8.1	200421
	Yermak Plateau	292.8	-40.5	19429	187107	23.3	251.6	200619
	Fram Strait	352.1	-43.2	21821	258969	16.8	118.1	200727
	North Pole - Lomonosov Ridge	15.6	-43.1	718	9642	18.8	102.3	200906
Data set used	MOSAiC region	Mean NASC (m^2 km^{-2})	Mean TS (dB)	Mean biomass (kg km^{-2})	Mean number (ind. km^{-2})	Mean length (cm)	Mean weight (g)	Mean date
Full Data Set	Eastern Amundsen Basin	5.8	-47.3	146	14649	12	28.5	191106
	Central Amundsen Basin	3	-51.4	51	14349	8.2	7.2	200121
	Western Amundsen Basin	0.5	-53.8	8	3984	6.1	3.9	200312
	Gakkel Ridge	0.1	-55.3	3	988	5.5	2.4	200329
	Nansen Basin	0.6	-52.1	11	2664	7	8.2	200421
	Yermak Plateau	422.5	-40.7	22784	646557	22.6	238.3	200619
	Fram Strait	477.9	-43.4	20708	2206180	16.3	115.2	200727
	North Pole - Lomonosov Ridge	19	-43.1	736	24908	18.3	107.9	200906
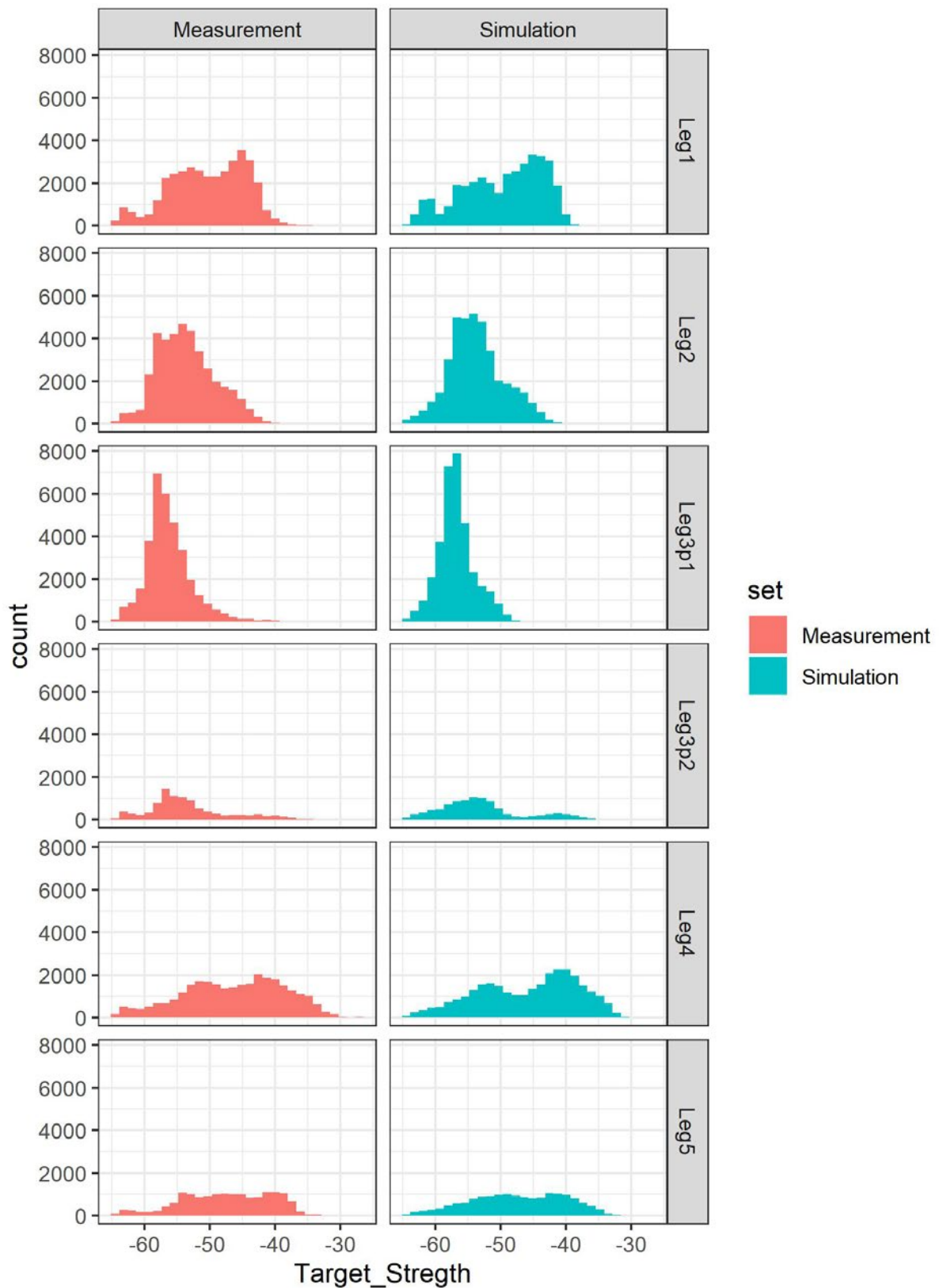
Figure 18: Results for MOSAiC Scenario 2: simulated TS distributions versus measured data. The left panel shows the actual measurements during the different legs of the MOSAiC expedition and the right panel shows the simulated fish community distribution based on the assumed combination of different taxonomic groups and different mean fish lengths in Scenario 2.
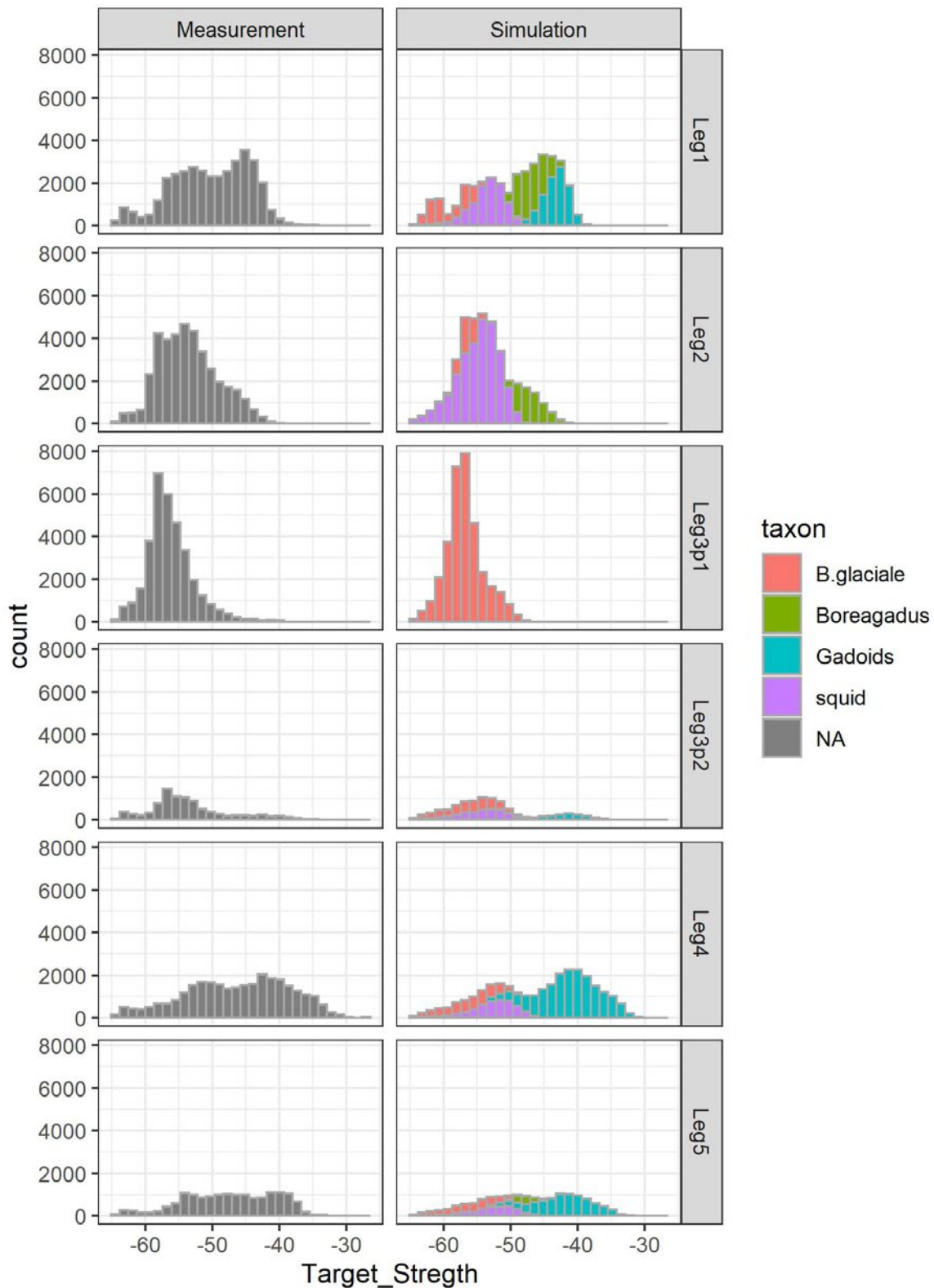
Figure 19: Results for MOSAiC Scenario 2: simulated TS distributions versus measured data. The left panels show the actual measurements during the different legs of the MOSAiC expedition. As there is no direct way to identify the species, the taxon id is indicated as "NA" (the grey colour). The right panels show the simulated fish community distribution based on the assumed combination of different taxonomic groups and different mean fish lengths in Scenario 2. These are the same data as in Figure 18 but the community distribution is broken down into taxonomic groups indicated by different colours.

MOSAiC: depth distributions of temperature and salinity along the expedition routes

One of the most prominent features of the Arctic Ocean is the circulation of subsurface warm and salty Atlantic water, here defined as the Arctic Atlantic water (AAW) and the dense Arctic Atlantic Water (dAAW) following Rudels et al. (2008)[64]. This water mass enters the Arctic Ocean through Fram Strait and the Barents Sea and circulates in the Arctic Ocean cyclonically, with one part branching off near the Lomonosov Ridge, following the ridge in the Amundsen Basin back towards Fram Strait. The Atlantic Water is continuously losing heat to the colder and fresher Arctic Surface Water (ASW) along its path.
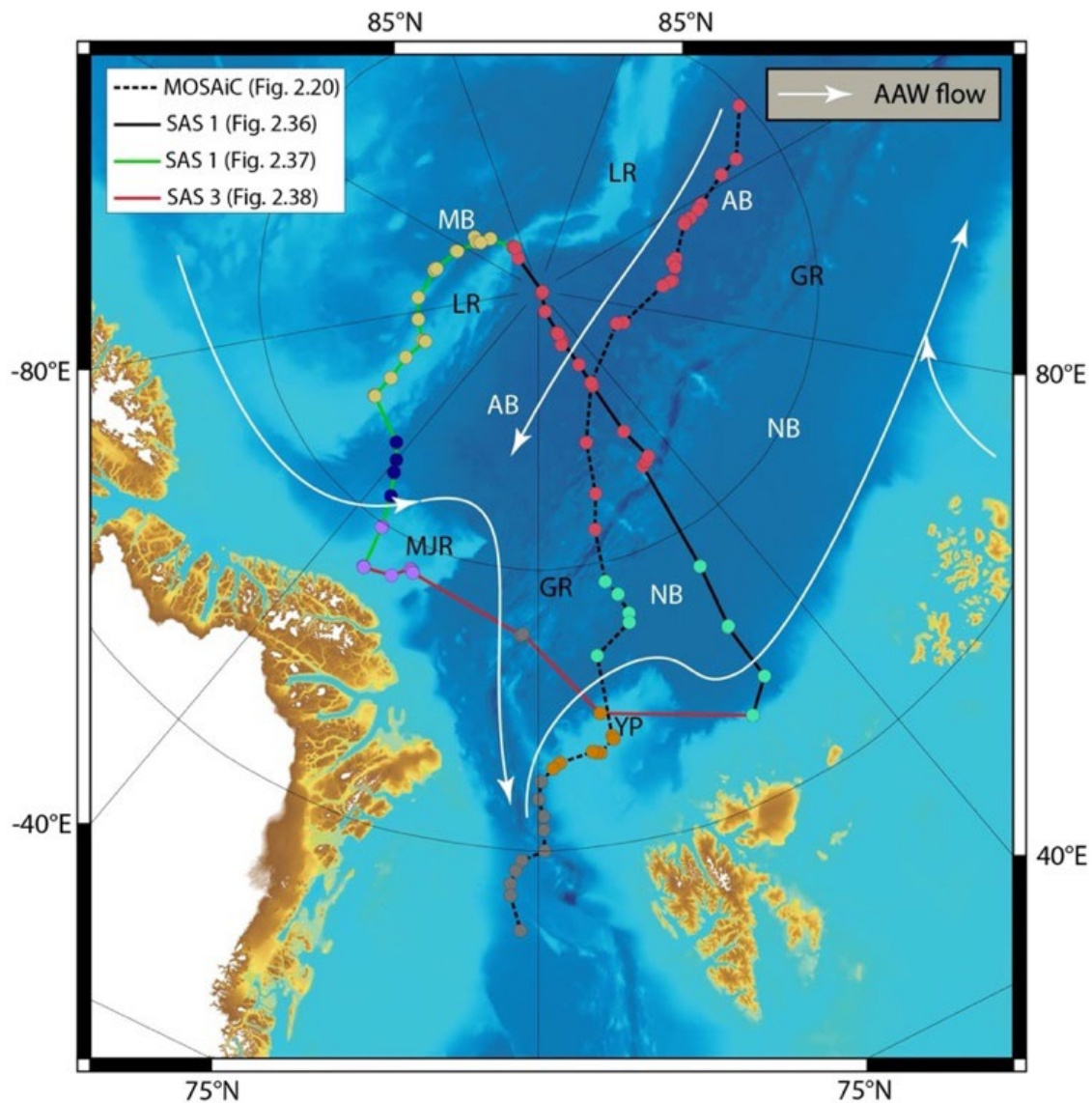
Figure 20: Map showing the CTD stations of the MOSAiC and SAS-Oden expeditions. Marked out on the map are the Nansen Basin (NB), Amundsen Basin (AB), Makarov Basin (MB), the Gakkel Ridge (GR), Lomonosov Ridge (LR), Morris Jesup Rise (MJR) and the Yermak Plateau (YP), indicated by different station colours. Also shown is the idealised Arctic Atlantic Water (AAW) flow pathways in the Arctic Ocean (modified from Rudels, 2012[65]).

---

[64] Rudels B, et al. (2008) Constraints on estimating mass, heat and freshwater transports in the Arctic Ocean: An exercise. In: Dickson RR, et al. (eds) Arctic-Subarctic Ocean Fluxes. Springer, Dordrecht [https://doi.org/10.1007/978-1-4020-6774-7_14]

[65] Rudels, B (2012) Arctic Ocean circulation and variability – Advection and external forcing encounter constraints and local processes. Ocean Science 8:261–286 [https://doi.org/10.5194/os-8-261-2012]

The first part of the MOSAiC CTD transect follows the return branch of the Atlantic Water that flows along the Lomonosov Ridge in the Amundsen Basin towards the Fram Strait. Along this first section of the transect (between 0 and 700 km), the gradual temperature decrease in the AAW is likely a combination of the continuous heat loss and the ship path deviating away from the AAW current towards the Gakkel Ridge. The transect then crosses over the Gakkel Ridge into the Nansen Basin, over the Yermak Plateau and finally into the Fram Strait area. The warm inflowing Fram Strait branch of the Atlantic Water is found in the transect from the Yermak Plateau and towards Fram Strait. The strong vertical salinity contrast between the cold and fresh ASW and the warmer and saltier AAW is often referred to as the cold halocline and is an ever-present feature in the Arctic Ocean. The dissolved oxygen concentration data from the MOSAiC CTD profiles show generally high concentrations (> 350 µmol kg^{-1}) above the cold halocline while below, within the Atlantic Water (AAW + dAAW) concentrations are fairly constant with depth and typically around 300 µmol kg^{-1} (MOSAiC expedition, unpublished data).

MOSAiC: fish distributions related to temperature, salinity and zooplankton abundance

Similar to previous oceanographic studies in the CAO, a slightly warmer and saltier Atlantic Water Layer (AWL) was observed at mesopelagic depths where the DSL was located, throughout the MOSAiC expedition. The trend in backscattering density closely followed the mean temperature spanning the layer between 200-600 m of depth. A similar trend also occurred in the TS distribution, which is indicative of target size, when larger targets were detected in the mesopelagic layer.

When NASC was plotted against salinity, it was evident that the larger targets and highest target densities resided in the AWL (MOSAiC expedition, unpublished data). Between Leg 1 and the end of Leg 3 the densities sharply declined where the coldest average temperatures were recorded in this section of transition between subpolar and polar water masses. Towards the end of Leg 3 and during Leg 4 when the drift route was in the southernmost sections of the expedition, the highest target densities were observed. It is important to note that, the oceanographic conditions in this section cannot be considered representative of the CAO anymore as the Atlantic water influence was strong during Leg 4. Higher biological productivity in this area and higher densities of Atlantic gadoids such as Atlantic cod and haddock is a known characteristics of this region and is referred to as the gateway into the Arctic because of the temporary suitability as preferred fish habitat and enhanced primary productivity accompanied with large year-to-year ice cover variability.

To predict both habitat suitability and potential density distribution for the fish biomass, a preliminary generalised additive modelling (GAM) approach was applied using the "mgcv" package in R[66]. This analysis is preliminary because the CTD data from the MOSAiC expedition have not yet been quality-assured and published. GAM is a nonparametric additive regression technique allowing for flexibility in exploring the potential associations between spatial patterns in explanatory variables such as temperature gradients, zooplankton availability, latitude, and fish backscatter density. Restricted maximum likelihood technique (REML) was used to determine optimal smoothing functions in the regression analyses as it avoids being constricted in local minima during the iterations (i.e., more accurate term predictions than Akaike Information Criterion).

The results of this preliminary GAM analysis (**Figure 21**) show that the four variables had high predictable power for the mean acoustic backscatter (NASC in m^2 km^{-2}) as 73.1% of the deviance was explained based on three environmental and biotic predictors: (1) average water temperature within the mesopelagic depth (200-600 m), (2) depth-integrated abundance of acoustically detected mesozooplankton abundance, (3) average salinity within the mesopelagic depth (200-600 m). The deviance explained

[66] Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73:3-36 [https://doi.org/10.1111/j.1467-9868.2010.00749.x]

is a similar indicator to the r-squared in the linear models. It measures the proportion of variation that the model accounts for, indicating the goodness of fit.
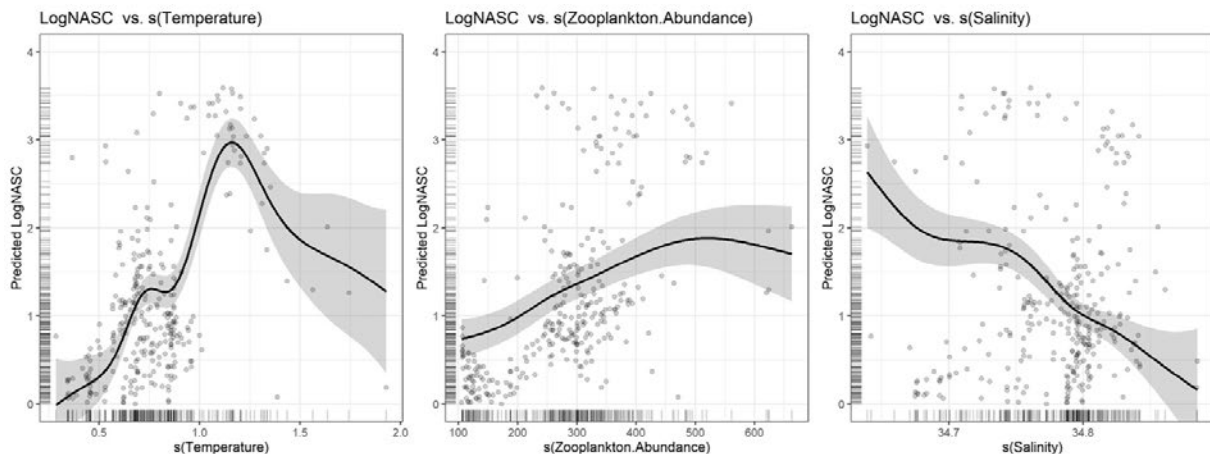
Figure 21: Results from the preliminary Generalised additive model (GAM) with fish acoustic backscatter (NASC m² nm⁻²) as response variable and three predictor variables to test if acoustically estimated fish density can be predicted based on environmental and biotic variables. Additive effects of the predictors on log-NASC. Left graph: Average water temperature within the mesopelagic depth (200-600 m). Middle graph: Depth-integrated abundance of acoustically detected mesozooplankton abundance. Right graph: Average salinity within the mesopelagic depth (200-600 m). The solid curves are smoothed functional responses, shading represent 95% confidence intervals.

SAS-Oden: target strength (TS) from single targets from EK80 data (18 kHz) and WBAT profiles (38 kHz)

Although the SAS-Oden expedition encountered significant noise interference from ice-breaking in certain areas, there were extensive stretches of high-quality data along the expedition route, allowing for the detection of individual targets with remarkable precision despite the low density of fish. The expedition's average TS distribution throughout its trajectory showed a relatively stable mean TS of approximately -45 dB (**Figure 22**). However, the highest TS values were observed along the Lomonosov Ridge in the western Amundsen basin, where the average target strength measured around -43.5 dB.

The most prominent TS peak in the EK80 measurements throughout the expedition was between -50 and -40 dB (**Figure 23**), suggesting dominance of *Boreogadus*. Targets suggesting significant presence of larger gadoids similar to MOSAiC expedition were never observed along the SAS-Oden expedition route except from the very few targets detected in the Yermak Plateau around -35dB. A secondary peak was observed between -60 and -50 dB mostly very weak, only pronounced in the Nansen Basin (**Figure 23**). These small targets consisted of only 31% of all detected fish tracks and could represent myctophids such as *Benthosema glaciale*.

The estimated length distributions based on the TS are shown in **Figure 24** for the EK80 and in **Figure 26** for the WBAT. The estimated size distribution was smaller from the measurements from the WBAT profiles especially in the second half of the expedition extending along the Morris Jesup Rise, Gakkel Ridge and Yermak Plateau, where a clear unimodal peak below 10 cm is visible. This peak potentially consists of mixture of juvenile polar cod and myctophid fish such as *Benthosema glaciale*. Thee estimated biomass for the mesopelagic zone resulted in comparable levels between the EK80 measurements and WBAT profiles (**Figure 25** and **Figure 26**). This confirms that, despite higher level of noise and being a relatively low frequency, the EK80 captured the distribution accurately.
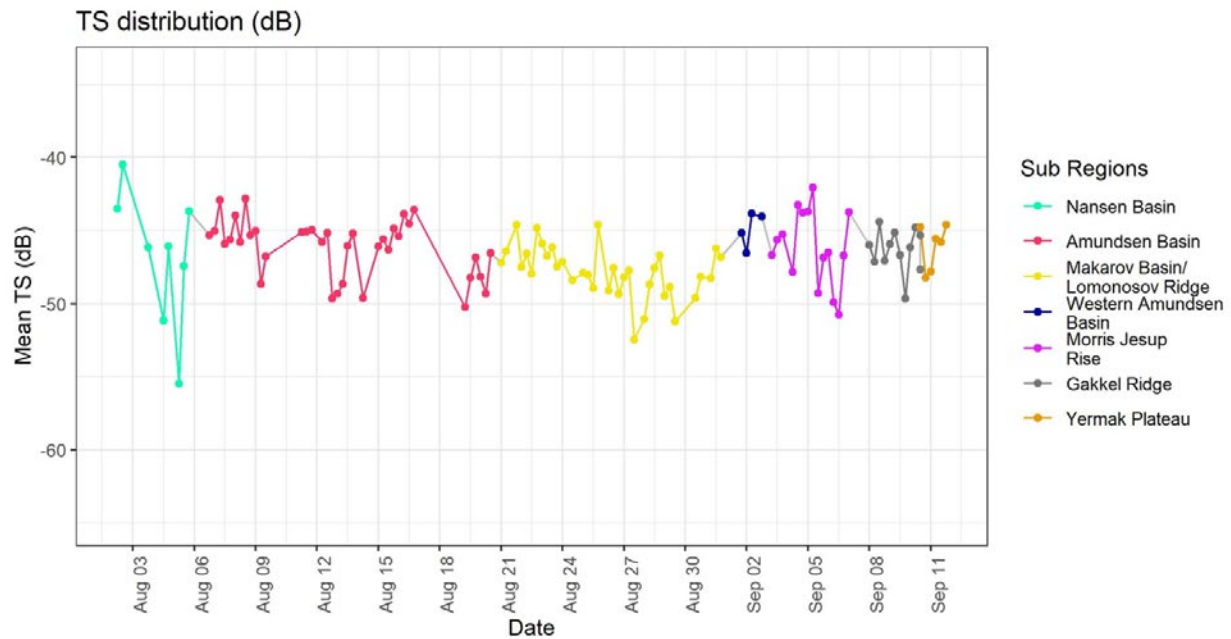
Figure 22: Target strength (TS) along the SAS-Oden route for the 20-600 m depth interval down to 600 m from the EK80. The mean TS shown here are from the fish-track detections averaged every six hours.

SAS-Oden: fish distributions, abundance and biomass from EK80 hydroacoustic data

The NASC measured during the SAS-Oden expedition had much lower intensity than that measured during the MOSAiC expedition. The backscatter levels along the entire SAS-Oden route are only comparable to those from the second half of MOSAiC Leg 2 and the first part of Leg 3 when the drift took place around the western Amundsen Basin, Gakkel ridge and Nansen basin. This is the section in the MOSAiC NASC data where the temperature of the mesopelagic zone drops to its lowest levels below 1 °C. In addition, the acoustic backscatter during the SAS-Oden expedition was not only characterised by the main dominance of the DSL in the mesopelagic layer, but there was also backscatter from the surface layer 0-100 m suggesting presence of *Boreogadus* in the surface layer (**Figure 27**).

Because of the high level of noise, the analysis method was slightly modified for final extraction of the acoustic densities. In the echogram sections where the targets could not be detected automatically due to noise, the good signals were manually boxed where the echo-traces showed clear fish patterns. These sections were then included in the SAS-Oden Target-tracking Data Set and there is no Full Data Set/Target-tracking Data Set separation is available for SAS-Oden. The mean NASC in different oceanic sections of the CAO within 200-600 m layer varied between 0.10-0.37 m^2 km^{-2} (in nautical terms equivalent to 0.34-1.27 m^2 nmi^{-2}), which is comparable to the Target-tracking Data Set from the Oden expedition in 2016[67] and slightly lower than their Full Data Set. In 2016, the mean NASC for the water column 20-600 m varied between 0.15-2.45 m^2 km^{-2} (0.51-8.40 m^2 nmi^{-2}), slightly larger than in the Full Data Set.

---

[67] Snoeijs-Leijonmalm P, et al. (2021) A deep scattering layer under the North Pole pack ice. Progress in Oceanography, 194:102560 [https://doi.org/10.1016/j.pocean.2021.102560]
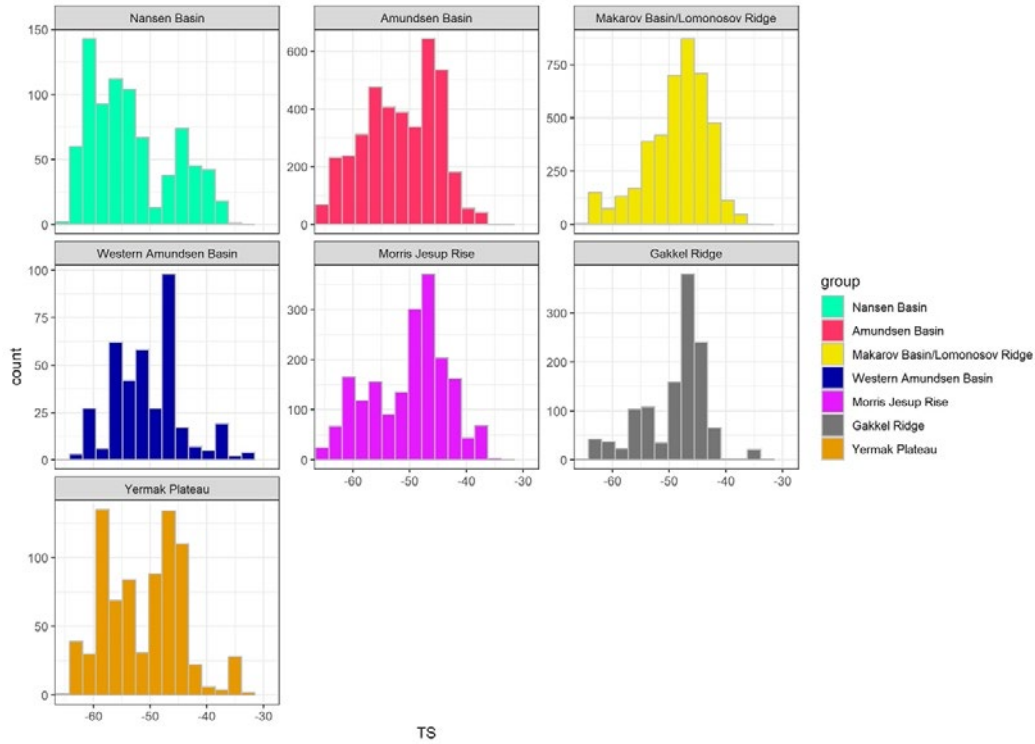
Figure 23: Distributions of the target strength (TS) for different sections along the SAS-Oden route from the EK80 measurements. The occurrence of multiple peaks are potentially due to mixture of different taxonomic groups such as the Arctic endemic gadoids *Boreogadus* and *Arctogadus* and small mesopelagic myctophids such as *Benthosema glaciale* for different sections SAS-Oden route. Here only ship mounted EK80 (18kHz) measurements are provided.
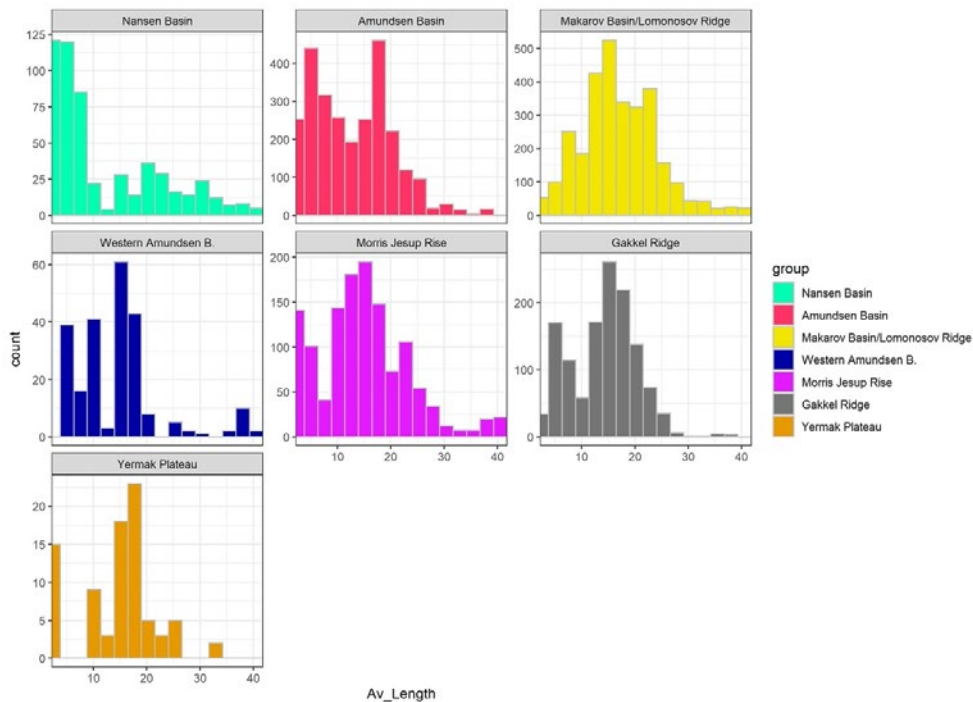
Figure 24: Distributions of estimated fish length (in cm) based on target strength (TS) and assuming dominance of *Boreogadus* and *Arctogadus* accompanied by the myctophid *Benthosema glaciale* for different sections SAS-Oden route from the EK80 measurements.
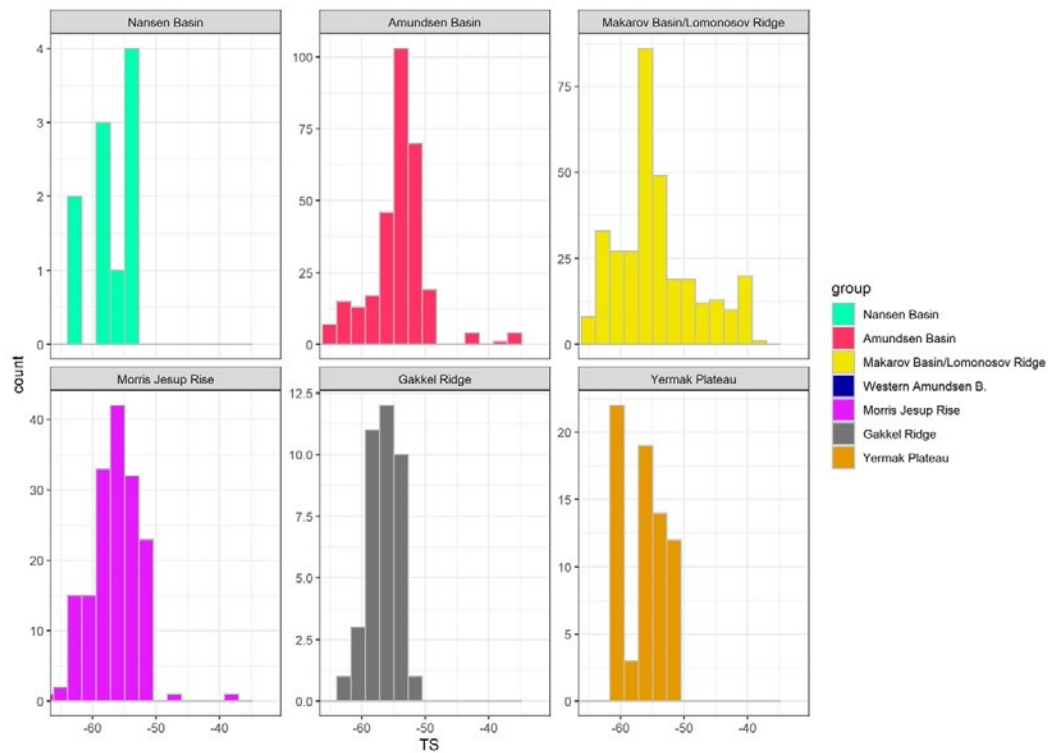
Figure 25: Distributions of the target strength (TS) for different sections along the SAS-Oden route from WBAT (38 KHz) profiles. The occurrence of multiple peaks are potentially due to mixture of different taxonomic groups such as the Arctic endemic gadoids *Boreogadus* and *Arctogadus* and small mesopelagic myctophids such as *Benthosema glaciale* for different sections SAS-Oden route.
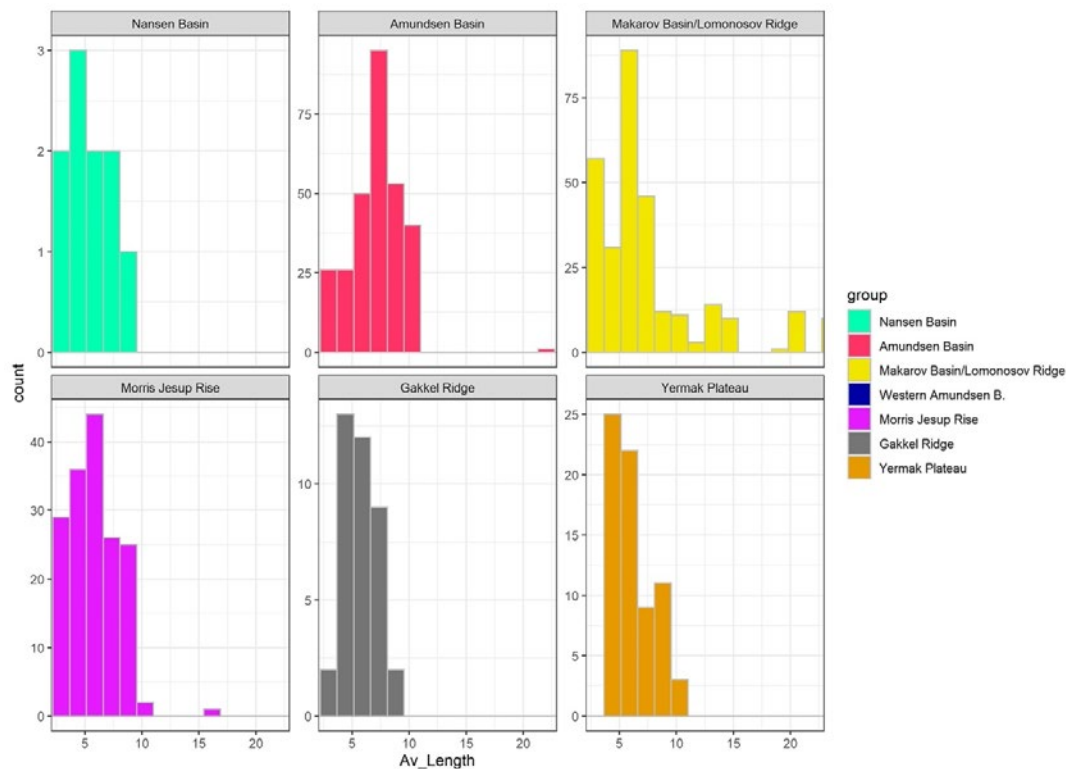
Figure 26: Distributions of estimated fish length (in cm) based on target strength (TS) and assuming dominance of *Boreogadus* and *Arctogadus* accompanied by the myctophid *Benthosema glaciale* for different sections SAS-Oden route from the WBAT measurements.

It is important to note that, the noise levels in the EK80 data were significantly lower during the SAS-Oden expedition in 2021 than during the Oden 2016 expedition, and some of the good signals in the water layer 0-200 m in 2016 must have been blocked by the noise and therefore were excluded from the analysis. The highest fish densities observed during the Oden 2016 expedition were close to the surface around the North pole and around Makarov basin/Lomonosov Ridge. In the upper colder portion of the water column (below 0°C), such higher densities most likely resulted from a mixture of juvenile and adult *Boreogadus*, but no density estimation was possible because of the noise.

Due to elevated surface backscatter around the Makarov Basin/Lomonosov Ridge, the NASC values in one location of this section reached up to 45.5 m^2 km^{-2} (156.1 m^2 nmi^{-2}), constituting the strongest backscatter observed throughout the expedition. Nevertheless, these are still substantially low densities compared to sub-polar regions, e.g., around the Yermak plateau during the MOSAiC expedition. Both during the Oden 2016 and SAS-Oden expeditions the mean NASC was highest at the stations closer to the North Pole and lowest around the western Amundsen basin and the Gakkel Ridge.

The geographical distribution of NASC densities during the SAS-Oden expedition are shown in **Figure 28**, and the average NASC and TS in different sections of the expedition are provided in **Table 21** and **Table 22**. The highest NASC values were observed in the upper part of the water column and north of 88 °N, i.e., around the North Pole. The density distribution in the mesopelagic zone was more uniform along the SAS-Oden route, with the lowest NASC values were around the Gakkel ridge and in the Amundsen basin, similar to the MOSAiC expedition. Biomass estimates based on these data are shown in **Figure 29** and **Table 21**.

Figure 27 shows the distribution of the backscatter throughout the expedition. At the start of the expedition, around the close-shelf sections of the southern Nansen basin, a relatively stronger mesopelagic layer is visible. As shown previously in **Figure 22**, in this section there are few relatively strong targets where the six-hour TS average goes up to -40 dB, and containing individual TS up to -35 dB suggests presence of relatively larger gadoids around 40 cm. However, after departure from the near shelf area, presence of such larger targets and higher density in the mesopelagic layer ceased to exist and a faint layer between 300 and 600 m remains further north along the expedition route.

A surface layer also appears in the Nansen basin between 40-80 m of depth (**Figure 27**), but disappears in the further Northern sections of the Nansen basin. A surface layer starts to appear again after reaching the Gakkel Ridge, positioned between 40 and 120 m with changing densities. This layer gradually becomes stronger towards the norther part of the Western Amundsen Basin continue to increase in density after the North Pole along the Lomonosov ridge. In Lomonosov ridge, this layer reaches to the highest density, also the TS measurements between the surface and 150 m reaches to maximum in this section (ca. -45 dB) Suggesting an average size of 18 cm while the third quantile is reaching to 25 cm. Further west in the Makarov basin this layer near the surface disappeared again. Looking at evolution of this surface layer from the edge to the center, and considering the polar cod as most likely candidate, it could be speculated that there is a North-ward movement in the polar cod distribution where larger/older fish are distributed close to the North pole and smaller-sized juveniles are found in the Nansen basin and southern parts of the Gakkel Ridge.

During the rest of the expedition, around the western Amundsen basin, Morris Jesup Rise and Gakkel Ridge, the densities remained extremely low with faint occurrence of the mesopelagic layer until the Yermak Plateau. A slight increase in the backscatter levels observed again in the Yermak plateau, however in densities far from being comparable to the weakest recordings by the Polarstern in this area a year ago.
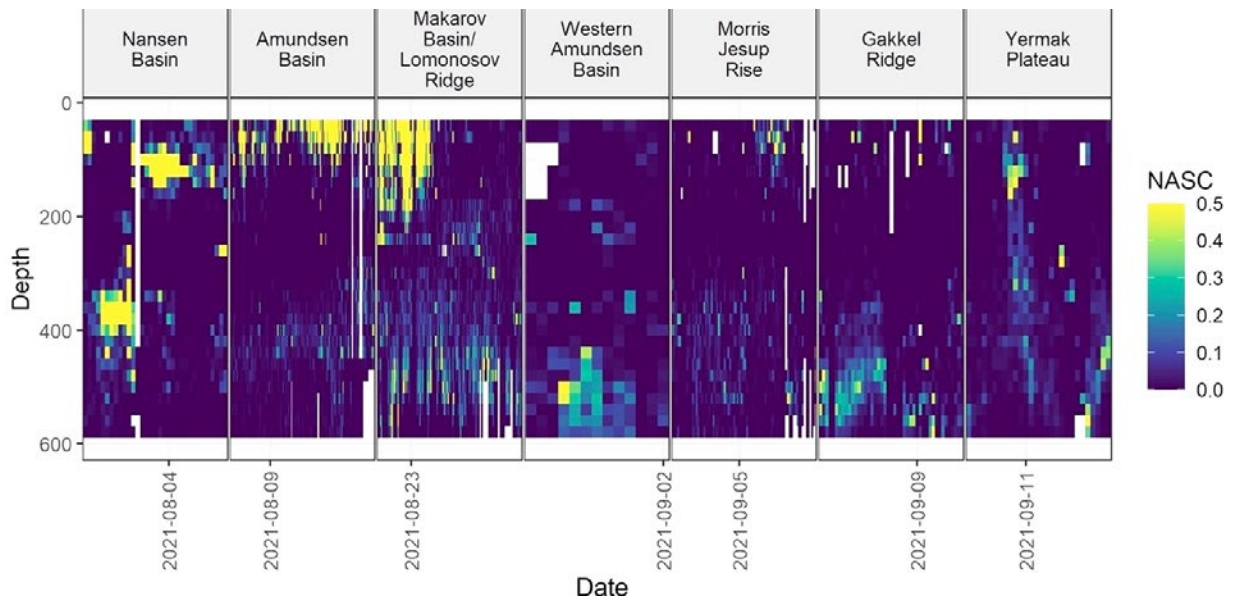
Figure 27: The Nautical Area Scattering Coefficient (NASC) measured along the SAS-Oden route. In contrast to the MOSAiC expedition, higher NASC values were observed in the upper part of the water column during the first half of the expedition, but still with comparatively low densities.
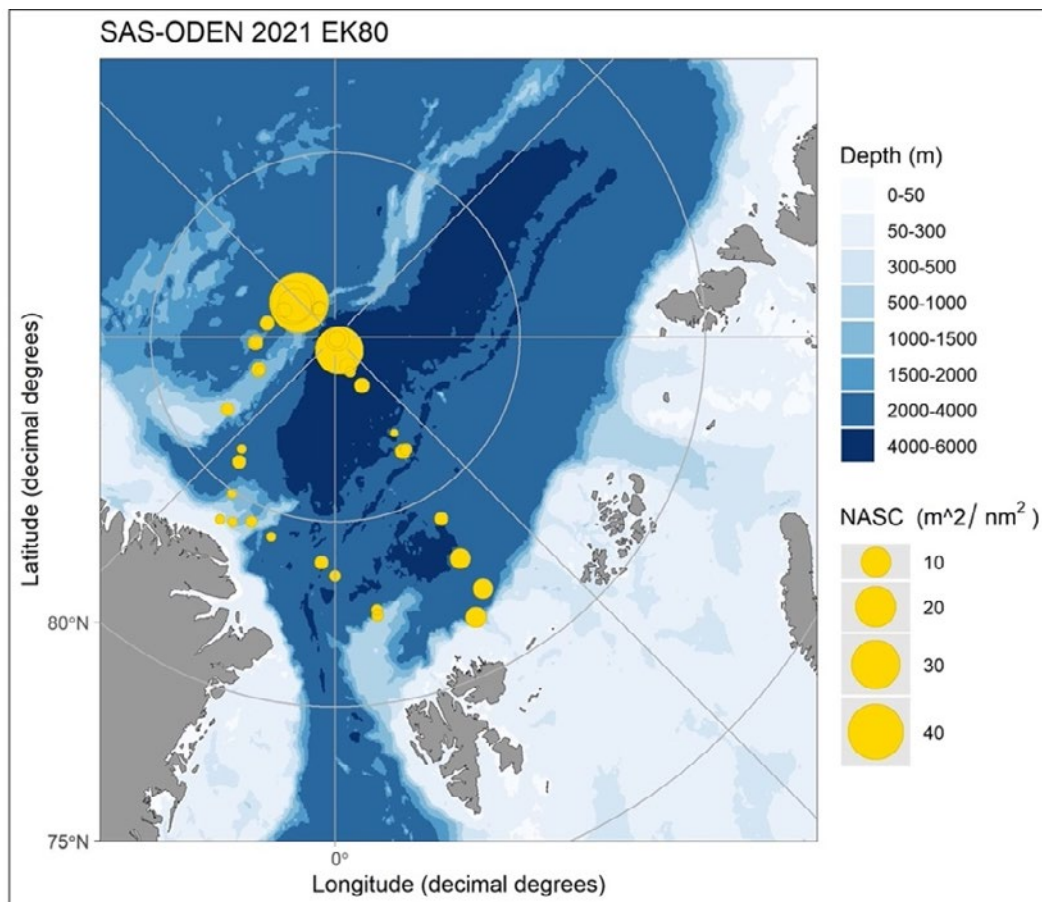
Figure 28: The Nautical Area Scattering Coefficient (NASC) distribution measured by the EK80, illustrating the density of fish-like targets along the SAS-Oden 2021 route at 20-600 m of depth based on EK80 measurements.
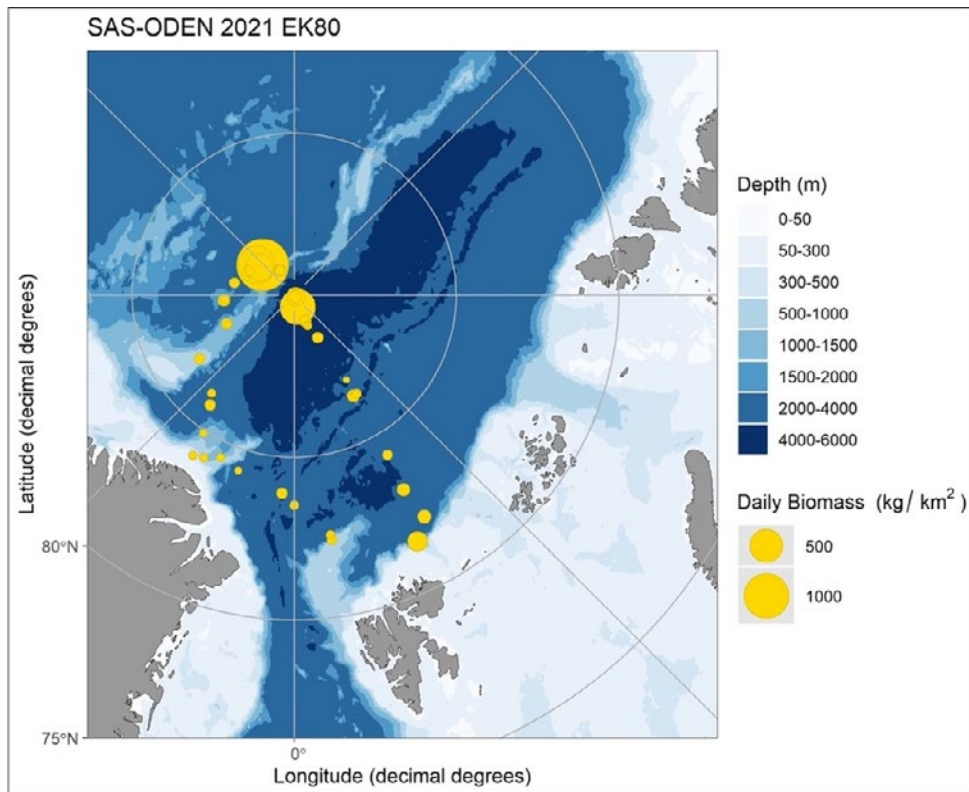
Figure 29: Estimated biomass distribution from the EK80 along the SAS-Oden expedition route taking the individually tracked targets into account for the size distribution, assuming *Boreogadus* and *Arctogadus* dominance.

Table 21: Average Nautical Acoustic Scattering Coefficient (NASC) measured by the EK80, estimated fish biomass, and estimated fish density for different sections of the SAS-Oden expedition in the full water column (20-600 m) and in the mesopelagic zone (200-600 m).

Region	Measured average NASC (m² km⁻²)	Estimated average fish biomass (kg km⁻²)	Estimated average fish density (number km⁻²)
Full water column (10-600 m)			
Nansen Basin	0.84	70	15666
Central Amundsen Basin	0.73	50	6588
Makarov Basin/Lomonosov	2.45	230	8836
Western Amundsen Basin	0.28	20	752
Morris Jesup Rise	0.15	10	756
Gakkel Ridge	0.25	20	1254
Yermak Plateau	0.26	20	1609
Mesopelagic zone (200-600 m)			
Nansen Basin	0.37	60	324
Central Amundsen Basin	0.10	10	235
Makarov Basin/Lomonosov Ridge	0.34	20	562
Western Amundsen Basin	0.26	20	268
Morris Jesup Rise	0.10	10	229
Gakkel Ridge	0.20	20	408
Yermak Plateau	0.22	20	681

Table 22: Average target strength (TS), estimated average fish length, and estimated average fish weight estimated from the EK80 data, as well as average temperature and average salinity for different sections of the SAS-Oden expedition in the full water column (10-600 m) and in the mesopelagic zone (200-600 m).

Region	Measured average TS (dB)	Estimated average fish length (cm)	Estimated average fish weight (g)	Measured average Temperature (°C)	Average salinity (PSU)
Full water column (10-600 m)					
Nansen Basin	-42.6	20.8	109	1.34	35.04
Central Amundsen Basin	-45.0	16.9	49	0.61	34.91
Makarov Basin/Lomonosov	-44.3	17.7	64	0.22	34.72
Western Amundsen Basin	-42.2	19.4	156	0.47	34.80
Morris Jesup Rise	-45.0	15.6	57	0.18	34.69
Gakkel Ridge	-46.6	13.6	33	0.67	34.95
Yermak Plateau	-45.4	14.3	56	0.94	35.01
Mesopelagic zone (200-600 m)					
Nansen Basin	-41.2	26.2	153	1.35	35.08
Central Amundsen Basin	-44.0	19.8	64	0.97	35.05
Makarov Basin/Lomonosov Ridge	-43.1	21.2	85	0.62	35.02
Western Amundsen Basin	-40.6	25.0	229	0.85	35.03
Morris Jesup Rise	-43.7	19.1	78	0.55	35.00
Gakkel Ridge	-45.3	16.4	45	0.93	35.05
Yermak Plateau	-44.1	17.6	77	0.91	35.06

SAS-Oden: depth distributions of temperature and salinity along the expedition route

The first CTD transect of the SAS-Oden expedition covered the Nansen and Amundsen basins through the North Pole to the Lomonosov Ridge (**Figure 20**). Thus, it covered the inflowing Fram Strait branch of the Atlantic Water in the Nansen Basin indicative of elevated temperatures in the AAW between 0 and 400 km and the outgoing branch towards the Fram Strait in the Amundsen Basin from about 800 km and onward. In the Nansen Basin, the cold halocline, separating the PSW and AAW is shallower and thickens gradually towards the North Pole. The dissolved oxygen concentration was generally high in the PSW (>350 µmol kg^{-1} (SAS-Oden expedition, unpublished data). The oxygen concentrations within the AAW were elevated in the southern Nansen Basin, associated with the core of the inflowing Atlantic Water mass, and decreases northwards. In the Amundsen Basin, from the Gakkel Ridge to the Lomonosov Ridge, oxygen concentrations within the AAW were somewhat patchy with regions with a clear oxygen minimum zone in the upper half of the AAW.

The second CTD transect of the SAS-Oden expedition (**Figure 20**) starts at the Lomonosov Ridge and passes southwards within the Makarov Basin and along the Lomonosov Ridge and into the western-most part of the Amundsen Basin, finally ending on the Morris Jesup Rise. The outer part of the Amundsen Basin return branch of the Atlantic Water was in the very first part of the transect (between 0 and 50 km) and again, perhaps as a mixture with the branch coming from the Canada Basin (it is not entirely clear where the branches merge) in the Amundsen Basin (between 650 and 750 km). The cold halocline was significantly deeper along this transect compared to SAS-Oden Transect 1, and salinities within the Atlantic Water somewhat lower (SAS-Oden expedition, unpublished data). Dissolved oxygen concentrations were significantly lower within the AAW in this transect compared to SAS-Oden transect 1 and to the MOSAiC transect. Just above the two Atlantic Water branches (between 0 and 50 km and between 650 and 750 km), inside the cold halocline there were prominent oxygen maxima. The AAW was generally associated with an oxygen minimum zone.

The third CTD transect of the SAS-Oden data set (**Figure 20**) started on the Morris Jesup Rise and went zonally over the Gakkel Ridge, separating the Amundsen and Nansen Basins, to the Yermak Plateau, and then to the starting point of SAS-Oden Transect 1 in the southernmost part of the Amundsen Basin. Note that the time difference between the Yermak Plateau station (last station of the SAS-Oden expedition) and the first station of the SAS-Oden expedition was about 7 weeks. This transect nicely showed the temperature difference between the outflowing and inflowing branches of the Fram Strait Atlantic Water as the corresponding difference in heat content has been lost to ice melt and to the atmosphere along its path. The outflowing branch is also associated with lower salinity and lower dissolved oxygen concentration.

SAS-Oden: fish distributions, abundance, and biomass from WBAT data

The WBAT profiles focused mainly on the DSL as the near-surface measurements were impacted by surface reflections. The DSL was resolved in high detail and the individual target tracks were clearly discernible. The average TS measured in the WBAT profiles was ca. 2.5 dB lower than the average TS from EK80 measurements. Therefore, the estimated average fish length was ca. 5 cm shorter compared with the EK80 measurements. On the other hand, the average NASC (**Figure 30**) and estimated biomass (**Figure 31**) were comparable between the two acoustic methods (compare **Table 21** and **Table 22** with **Table 23** and **Table 24**). Although it could be considered counter-intuitive to estimate similar level of biomass while the detected size distribution is smaller from the WBAT, in fact many small-sized fish would give similar biomass as the few large fish. Eventually, the NASC level determines the biomass which was comparable between the two systems. The reason that the number of detected larger targets are too small with the WBAT could be potentially due to very short deployment time when compared to the continuous measurements from the EK80.

Other reasons for smaller estimated sizes could be due either to higher sensitivity of the WBAT as it was substantially less susceptible to noise, or potential avoidance of the CTD with WBAT by the targets. It is well-known that mesopelagic nekton communities escape from the lowered equipment under water, and this also hinders the capability of representative biological samples by the trawl nets[68]. Active avoidance of submerged equipment by fish was also observed on the EK80 during the SAS-Oden expedition. Nevertheless, for the WBAT such avoidance is probably not a big issue because the WBAT had a range of 200 m under the transducer. Thus, target detection would start before avoidance reactions. The WBAT measurements agreed with the EK80 measurements during most of the SAS-Oden expedition, but on the Yermak Plateau the WBAT NASC values were higher those of the EK80 measurements. This difference might have been caused be higher microzooplankton density at the Yermak Plateau, especially considering large zooplankton individuals.

Comparison of results

From the acoustic backscatter it can be inferred that changes in the geographical distribution of both abundance and size are in response to vertical and horizontal temperature gradients, seasonal/permanent stratification, zooplankton density, and seasonal factors. The interplay between these factors is complex and may also be influenced by other factors that were not directly accounted for in this chapter, including ice coverage, light intensity, primary productivity, and bio-geo-chemical dynamics.

---

[68] Kaartvedt S, et al. (2012) Efficient trawl avoidance by mesopelagic fishes causes large underestimation of their biomass. Marine Ecology Progress Series 456:1-6 [https://doi.org/10.3354/meps09785]

Table 23: Average Nautical Acoustic Scattering Coefficient (NASC) measured by the WBAT, estimated fish biomass, and estimated fish density for different sections of the SAS-Oden expedition in the mesopelagic zone (200-600 m).

Region	Measured average NASC (m² km⁻²)	Estimated average fish biomass (kg km⁻²)	Estimated average fish density (number km⁻²)
Mesopelagic zone (200-600 m)			
Central Amundsen Basin	0.73	44.0	4434
Makarov Basin/Lomonosov	0.65	35.8	2450
Morris Jesup Rise	0.24	14.6	344
Gakkel Ridge	0.47	33.3	1370
Yermak Plateau	1.95	123.2	22850

Table 24: Average target strength (TS), estimated average fish length, estimated average fish weight, average temperature and average salinity for different sections of the SAS-Oden expedition in the mesopelagic zone (200-600 m).

Region	Measured average TS (dB)	Estimated average fish length (cm)	Estimated average fish weight (g)	Measured average Temperature (°C)	Measured average salinity (PSU)
Mesopelagic zone (200-600 m)					
Central Amundsen Basin	-45.6	16.2	42.6	0.70	34.91
Makarov Basin/Lomonosov	-45.9	15.8	42.8	0.27	34.80
Morris Jesup	-44.1	19.0	75.2	0.30	34.83
Gakkel Ridge	-47.4	12.6	23.3	0.43	34.91
Yermak Plateau	-49.9	9.7	10.4	1.16	35.04
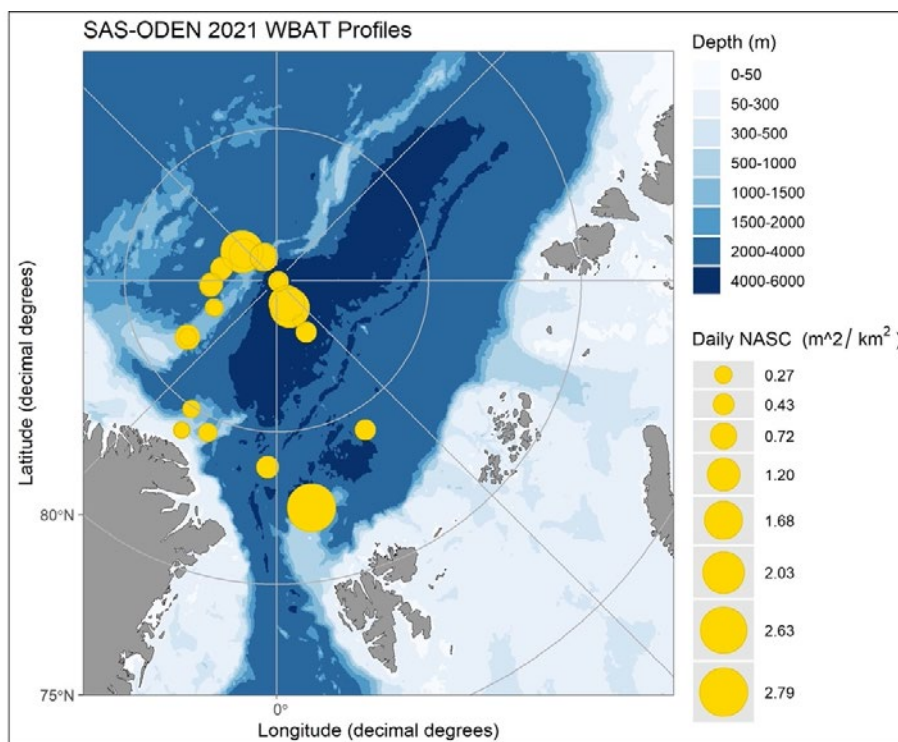
Figure 30: The Nautical Area Scattering Coefficient (NASC) distribution measured by the WBAT, illustrating the density of fish-like targets along the SAS-Oden 2021 route at 200-600 m of depth.
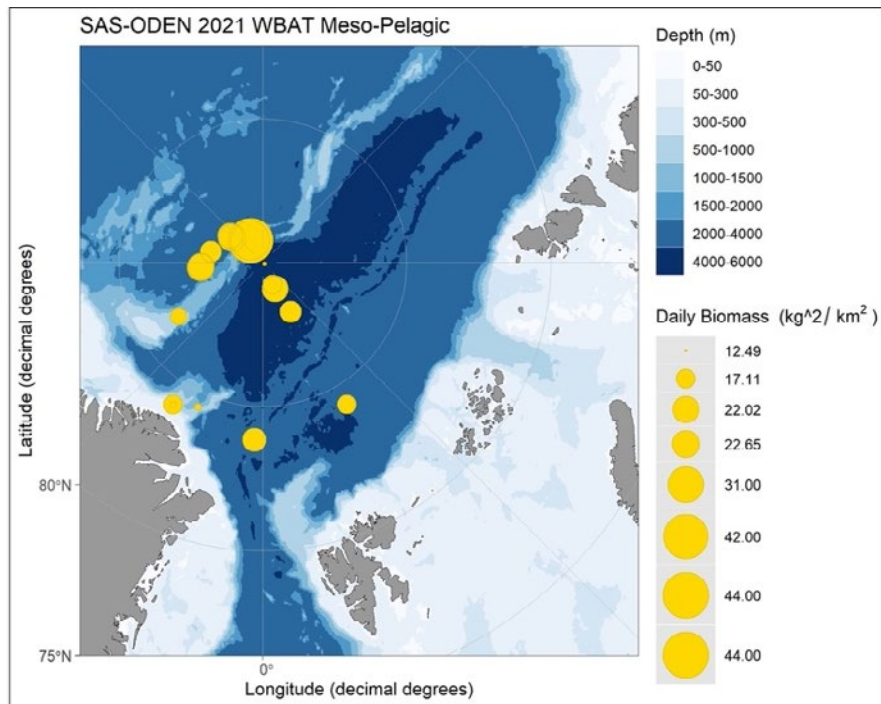
Figure 31: Geographic distribution of estimated fish biomass based on the acoustic backscatter measured by the WBAT during the SAS-Oden expedition corresponding to the same section as the NASC in Figure 28 (200-600 m). The biomass estimation was performed assuming *Boreogadus saida* as the dominant taxon with a significant contribution from the mesopelagic myctophid *Benthosema glaciale*.

Using the data collected during the MOSAiC expedition, a preliminary GAM model was developed to establish the relationship between fish density and explanatory variables. The model was able to predict the distribution of fish density with a high degree of accuracy, taking into account the nonlinearities in the relationships. During the expedition, high fish densities were observed in the Fram Strait and the Yermak Plateau. These areas are characterised by the marginal ice zone and close proximity to the continental shelf, which are associated with high biological productivity. This is in contrast to the CAO, which has a lower biomass carrying capacity. It is known that very high densities of gadoids, such as Atlantic cod and haddock, reside in the Sub-Arctic Norwegian and Barents Seas and migrate into seasonally ice-covered waters, with an increasing trend as a result of climate change impacts. These dynamic areas with large year-to-year ice cover variability and primary productivity coincide with the Atlantic water inflow to the CAO.

The findings indicate that biomass densities in the gateway regions were higher than previously reported. The Yermak Plateau had an average biomass of 19 tonnes km^{-2} and Fram Strait had an average of 22 tonnes km^{-2}, with local daily averages reaching up to 308 and 322 tonnes km^{-2}, respectively. Additionally, this study reveals the northward extent to which fish from the North Atlantic can migrate under the ice, providing insight into how the Arctic ecosystem may respond to ongoing climate change. However, the dynamic and complex nature of these habitats makes it challenging to fully characterise these regions and predict future changes. Furthermore, such high-density patterns were not observed during the SAS-Oden expedition in the same region, which is not surprising given the expected patchy distribution of a seasonally migrating pelagic/mesopelagic fish community.

Based on these observations, it can be concluded that environmental factors alone cannot accurately predict the precise locations of fish biomass within their preferred habitats, as stochasticity in geographical positioning and seasonal phenomena can play a significant role in determining their distribution. Therefore, a regional assessment of total

biomass can only be produced through a fishery survey that utilises systematic transects to account for the randomness in distribution. Although executing such surveys in ice-covered regions can be challenging, the findings of this study can be used as a guide for designing mini-surveys focused on relatively narrow regions that are highly representative of larger regions with similar features.

This study's findings can be used to determine the necessary baseline information for future fishery surveys in ice-covered regions, including survey season, stratification, transect design, total required effort, and equipment. Based on the GAM analysis, potential high-density regions can be identified as hot spots for transect densities and sampling efforts. The results not only confirmed the high-density regions in the marginal ice zone but also demonstrated the extremely low densities further into the CAO. The sudden drop in biomass observed immediately north of the gateway zone in the southern Nansen basin highlights the dominance of cold Arctic water in the first 100-150 m, which is virtually devoid of life. This confirms the findings of several previous studies that only detected faint signals of low fish biomass in a relatively warmer mesopelagic zone at 200-600 m.

In the mesopelagic section influenced by Atlantic water intrusion, the presence of small fish such as *Benthosema glaciale* was predominant. However, polar cod was expected to dominate the fish fauna in this region due to its known circumpolar distribution, except for a small region in the southern Nansen Basin during the SAS-Oden expedition. The lack of prominent densities of polar-cod-like backscatter between 84°N and 88°N makes this section of the Eurasian basin the poorest region in terms of fish density. Surprisingly, increased densities of polar-cod-like targets were observed during both the MOSAiC and SAS-Oden expeditions in the regions near the North Pole, Lomonosov Ridge, and the eastern (Siberian) side of the Amundsen Basin.

The observed differences in fish densities across various geographical locations under permanent ice cover are a crucial finding of this study as they emphasise the significance of gaining a better understanding of how the circulation of water masses under the ice can affect the biological productivity of the mesopelagic zone in the CAO, with distinct pathways and local biological dynamics. This knowledge could prove valuable for future projections. Although these high latitudes are expected to remain ice-covered in the coming decades, the conditions underneath the ice could be affected by changes in stratification and circulation patterns.

2.6. Answers to the WP2 research questions

(1) What is the potential fish abundance along the MOSAiC and SAS-Oden expedition routes as estimated from acoustic data?

Along the routes of the MOSAiC and SAS-Oden expeditions, acoustic tracks of individual targets indicating various size groups of fish were detected, and changing densities of acoustic backscatter were recorded. In the CAO the acoustic signals were mostly detected as weak reflections from the central Arctic DSL (200-600 m) in the mesopelagic zone. Evidence of elevated backscatter north of 88 °N during both the MOSAiC and SAS-Oden expeditions, although in different ways. The SAS-Oden observations around the North Pole and Makarov Basin/Lomonosov Ridge section were mainly characterised by surface scatterers. During MOSAiC increased densities were mainly associated with the mesopelagic zone below 200 m and indications of the presence of relatively larger fish were found east of the North Pole in the Amundsen basin and slightly west of the North Pole near the Lomonosov Ridge.

Target strength (TS) distributions were diverse and varied with depth and location. Size estimations from the detected fish individuals often showed multi-modal distributions suggesting assemblages of different life stages and/or different species. This required assumptions regarding potential constituents of these assemblages from the small

available evidence of taxon composition. When assuming dominance of the Arctic endemic gadoids *Boreogadus saida* and *Arctogadus glacialis* ("MOSAiC Scenario 1") the range of fish density varied between 4 000 individuals km^{-2} (southwestern Amundsen basin and Gakkel Ridge) to 450 000 individuals km^{-2} near the Yermak Plateau at the southern edge of the CAO in the Atlantic water inflow to the CAO. Strangely enough, such high fish abundances were not found in the latter area during the SAS-Oden expedition, and it may be hypothesised that inflow of fish through the Atlantic Gateway is a pulse-like, perhaps seasonal, phenomenon.

When a more complex taxonomic composition was assumed, based on the few catches and video observations in the CAO (see WP3 and WP5), with presence of Arctic gadoids and Atlantic gadoids of various size groups as well as the small mesopelagic myctophid *Benthosema glaciale* and armhook squid *Gonatus fabricii* ("MOSAiC Scenario 2"), the range of fish density varied between 6 400 individuals km^{-2} (southwestern Amundsen basin and Gakkel Ridge) to 2 570 000 individuals km^{-2} near the Yermak Plateau at the southern edge of the CAO in the Atlantic water inflow to the CAO. During the SAS-Oden expedition fish density was always low, varying between 800 and 15 700 individuals km^{-2} in different sections of the expedition route, assuming dominance of the Arctic endemic gadoids accompanied by *Benthosema glaciale*.

(2) What is the potential fish biomass along the expedition routes of the MOSAiC and SAS-Oden expeditions as estimated from acoustic data?

Estimated biomass from the MOSAiC expedition ranged from 20 kg km^{-2} to 22 000 kg km^{-2} where the lower values were estimated from the boundary section between subpolar and polar water masses north and south of the Gakkel Ridge. The higher estimates are from the Yermak Plateau (outside the CAO) and the southern Nansen Basin along the sections where the MOSAiC drift expedition was close to the continental slope and the Atlantic water inflow to the CAO. According to "MOSAiC scenario 2" where a more complex fish assemblage was hypothesised, the biomass range changed from 10 kg km^{-2} to 181 000 km^{-2}. These numbers were estimated based on established length/weight relationships from the literature where the length distribution was estimated based on the TS distribution.

(3) How does fish abundance and biomass relate to zooplankton abundance and biomass? (collaboration with WP3)

The zooplankton biomass during the SAS-Oden expedition was generally higher in the Eurasian Basin, on the Yermak Plateau and on the Lomonosov Ridge than in the Makarov Basin and on the Morris Jesup Rise (see WP3). This general pattern was reflected in the distribution of hydroacoustic backscatter and fish biomass except for the Yermak Plateau. Remarkably, the area near the Lomonosov Ridge where estimated fish biomass was an order of magnitude higher than during the rest of the expedition was also characterised by relatively high zooplankton biomass. In conclusion, the comparison of the spatial distribution of fish biomass and zooplankton biomass suggests that fish distribution generally mirrors the distribution of their zooplankton prey in the CAO. However, statistically the relations between fish and zooplankton abundance and biomass were not significant.

(4) What is the physical and biotic environment in which the potential fish resided during the two expeditions?

The fish backscatter density was mainly driven by temperature where the higher densities were associated with temperatures above 1 °C. The results showed a clear trend in backscattering density closely following the mean temperature and the density declines when the average temperature of the water layer between 200 and 600 m goes

below 1 °C. Temperature not only seemed to control fish density but also the size of the organisms as there is a similarly noticeable trend in the TS distribution (which is indicative of the size of the targets) where larger targets were found in the warmer section of the meso-pelagic layer. When combined with the salinity, it is evident that the larger targets and the highest densities are residing in the Atlantic water layer at 300–600 m of depth. Between Leg 1 and end of Leg 3 of the MOSAiC expedition the densities sharply declined in the area where the coldest average temperatures were recorded in this section of transition between subpolar and polar water masses. Towards the end of Leg 3 (southward drift) and during Leg 4 when the drift route was located in the southernmost sections of the expedition, and the distance to the ice margin and continental shelf was smallest, the highest fish densities were observed. Based on such these strong relationships, a generalised additive model (GAM) was used to test and model whether acoustically estimated fish density can be predicted based on the based on available explanatory variables. The four variables, temperature, latitude, depth and zooplankton abundance had high predictable power for the backscatter from fish, explaining 57.7% of the deviance.

2.7. Relevance of the WP2 data for fish stock modelling

Stock assessment is a crucial tool for effective fisheries management, aimed at maintaining sustainable harvest levels that balance population health with economic gains. The process involves predicting future stock abundances under different management scenarios by analysing a broad range of parameters such as age structure, spawning patterns, mortality, and growth rates, which require continuous monitoring over long periods. Reliable projections can be made only when these parameters are recorded over many years.

In this chapter, a detailed distribution of fish abundance and biomass in various sub-regions of the Arctic Ocean is presented, estimating size distribution of targets through backscattering strength. The results of this study establish an essential baseline for future assessments. A preliminary generalised additive modelling approach was used to explore the influence of environmental variables, including temperature, salinity, and zooplankton abundance, on fish communities of both Arctic and Atlantic origin. The Arctic Ocean is highly susceptible to climate-related changes, which are expected to lead to shifts in fish stocks' distributions, presenting significant challenges for stock assessment and fisheries management practices.

The results of this work contain uncertainties due to a lack of representative biological samples, making it unsuitable for direct stock assessments at this stage. However, a realistic range of species, size, and biomass distribution is provided over a vast geographical area in the Arctic Ocean for the first time, serving as an important reference point for future studies.

2.8. Recommendations from WP2 for the JPSRM of the CAOFA

Define standard JPSRM data collection settings for ship-mounted echosounders. To get the best possible data the echosounder should be dedicated to fish and zooplankton targets. Transducers should be at least 38 kHz (standard for fish), 18 kHz (deeper signals possible) and 200 kHz (standard for mesozooplankton).

All noise disturbances of the acoustic data should be avoided as much as possible. There are many sources of noise that come from the ship and this should be tested thoroughly and alleviated before leaving for a long expedition to the CAO. Noise from ice-breaking is unavoidable, but it is possible to make regular stops, turn off the ship's engines and collect good acoustic data for 10 minutes (e.g., every hour). Drift expeditions are best for acoustic surveys (no ice-breaking, no engines on). Acoustic measurements away from the ship, e.g., WBAT by helicopter can also be a solution.

Recommended acoustic equipment on the CTD. Back-up equipment is mandatory – motherboards are sensitive and can usually not be repaired on board.
  * WBAT 200 or 333 kHz (zooplankton)
  * WBAT 38 kHz (fish, but only if no ship-mounted echosounder is available)
  * Setup of a very precise accelerometer and compass, or a CTD stabiliser preventing tilt and rotation

Develop a standard "JPSRM set-up" for on-board measurements of zooplankton acoustic properties. These measurements needs to be performed when the zooplankton organisms are alive. Accurate knowledge on macrozooplankton abundance is important for predicting the potential preferred fish habitat.

Deploy "JPSRM ice-tethered autonomous acoustic buoys" that can transmit data to land. This is a relatively cost-effective way to get a lot of acoustic data[69].

---

[69] Flores H, et al. (on-line manuscript) Sea-ice decline makes zooplankton stay deeper for longer, 09 January 2023, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-2436026/v1]

3. IN SITU OPTICS AND ZOOPLANKTON ACOUSTICS (WP3)

3.1. Research questions addressed by WP3

(1) When and where did nekton (fish, squid) occur along the MOSAiC and SAS-Oden expedition routes?
(2) How is the water column structured in terms of particle distribution as a measure of ecosystem productivity?
(3) How is the water column structured in terms of zooplankton abundance as possible feeding grounds for fish?
(4) How does zooplankton abundance estimated by optical and hydroacoustic data relate to net data? (collaboration with WP4)
(5) How does fish abundance and biomass relate to zooplankton abundance and biomass? (collaboration with WP2)

3.2. Data produced by WP3

EFICA_DATA_SC07-WP3_A.UVP.xlsx
EFICA_DATA_SC07-WP3_B.LOKI.xlsx

For details of the Device Operations (date, time, geographical position, station depth), see files "MOSAiC_Device_operations" and "SAS-Oden_2021_Logbook"[70].

UVP6 and LOKI data from the MOSAiC expedition will become publicly available on PANGAEA[71] (these data are analysed within other MOSAiC projects than the SC07 project), and are not included in this chapter.

3.3. Human resources of WP3 and main responsibilities

Barbara Niehoff (AWI) coordination of the optical data analyses, MOSAiC FishCam automatised video analyses, UVP6 data analyses; Nicole Hildebrandt (AWI) LOKI data analyses; Pauline Snoeijs-Leijonmalm (SU) MOSAiC FishCam manual video analyses, SAS-Oden FishCam video analyses; Serdar Sakinan (WMR) Hydroacoustic analyses targeting zooplankton

3.4. Methods used by WP3

Analysis of the MOSAiC FishCam video recordings

Altogether, the video recordings from the MOSAiC FishCam consist of 1938 hours (ca. 3.9 TB)[72]. The data from 180 hours were already analysed manually for the occurrence of fish, squid and macrozooplankton. Nekton was usually observed within ca. 0 to 3 m from the camera objective, and zooplankton was observed within ca. 0.0 to 0.6 m distance to the camera, while the total visible distance from the camera lens was up to max. 10 m. The results from these 180 hours have been published[73]. Manual analyses are extremely time-consuming because the nekton is swimming extremely fast, in some cases only covering one or a few video frames. Therefore, the movies cannot be analysed at accelerated speed and, even then, it is easy to miss the nekton because they are moving

---

[70] Bolin Centre Database, Stockholm University [https://bolin.su.se/data/]
[71] PANGAEA Data Publisher for Earth & Environmental Science [https://www.pangaea.de]
[72] Snoeijs-Leijonmalm P, et al. (2021) Ecosystem mapping in the Central Arctic Ocean (CAO) during the MOSAiC expedition. Publications Office of the European Union [https://data.europa.eu/doi/10.2926/714618]
[73] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

so fast. Within WP3, an algorithm was developed by Simon-Martin Schröder (University of Kiel, Germany) to automatically detect a change (movement of organisms) compared with an empty background (no movement of organisms). Video frames (images), were then stored in a folder for manual identification of the cause of the change. With this method the complete set of video recordings was analysed yielding >250 000 images potentially presenting nekton.

The first 10 seconds of each video file were skipped because the camera adapted its brightness when the filming started. Also, the first 8 lines of each frame were cut off because they contained black pixels only. To estimate an empty background, the median grey value of each pixel over 21 frames was calculated and the resulting image was blurred. The median was updated for every 15th frame. The estimated background was then subtracted from each frame. A threshold was applied to obtain a mask of foreground objects and gaps smaller than 4 x 4 pixels (ca. 0.05 x 0.05 mm) were removed. Objects larger than 4 x 4 pixels were detected and tracked over time using an object tracker. The properties of each tracked object in each frame were stored in a "Comma Separated Values" (CSV) file.

Tracks of sharp objects longer than 0.5 second (15 frames) and shorter than 30 seconds with a minimum area of 64 pixels² were automatically selected. For each tracked object, the sharpest image (maximum variance of the Laplacian distribution, i.e., a continuous probability distribution named after Pierre-Simon Laplace) in a time interval of 1 second was selected. Multiple images were usually stored for one and the same tracked object. Large images were resized to a maximum of 512 x 512 px. The resulting images were fed into an "ImageNet-pretrained ResNet18 CNN Model" to extract 512-dimensional feature vectors. Images and feature vectors were processed using the "MorphoCluster" image annotation application[74] in which images are clustered by their optical properties using the "HDBSCAN procedure". Emerging clusters were validated manually, i.e., homogenous clusters with similar images were confirmed while inhomogeneous clusters were rejected, which puts the respective images back into the pool of un-clustered images. Then, the clusters were grown, i.e., MorphoCluster suggested images for each cluster from the un-clustered image pool, and these suggestions were either manually confirmed or rejected. When all clusters of one clustering process had been validated and grown, the clustering was started anew with the remaining un-clustered images. In total, this iterative three-step process of clustering, validation and growth was performed thirteen times until 65% of the data set was sorted into a total of 571 clusters of similar images. At that point, further clustering did not yield any new homogenous clusters. Therefore, the existing clusters were merged into categories and both, clustered and un-clustered images were uploaded to EcoTaxa (in total 225 820 images).

In EcoTaxa, the (previously clustered) images in each category were checked for consistency and, if necessary, manually corrected. Also, categories for the un-clustered images were predicted using the clustered images as training set, and manually evaluated. The taxonomic groups that could be identified beyond doubt were fish (61 images), cephalopods (344), chaetognaths, siphonophores, copepods, malacostraca crustaceans, of which most were amphipods, and ctenophores. The vast majority of the images, however, showed blurry objects that could not be allocated to any taxonomic category. Likely they included mostly copepods, hydromedusae, and amphipods as these groups were the most frequent on good images, and neuronal network approaches might allow to further analyse these images. This was, however, beyond the scope of the present study.

When all images were evaluated, tables were extracted for each category from EcoTaxa, which contained, among many other parameters, a tracking identification number that was created by the video analysis algorithm. In theory, images with the same 16-digit ending should have been the same organism, based on morphological criteria and back-

---

[74] Schröder SM, et al. (2020) MorphoCluster: Efficient annotation of plankton images by clustering. Sensors 220:3060 [https://www.mdpi.com/1424-8220/20/11/3060]

tracking of its appearance on the video frame. However, since the animals moved during filming, their appearance often changed tremendously, and thus the tracking ID yielded more organisms than there actually were. For fish and cephalopods, the images on EcoTaxa were re-analysed, and each observation was verified via its timestamp. This showed that the number of images for an individual cephalopod or fish, as extracted from the video, varied between 1 and 46, depending on the time that an individual spent in front of any of the two cameras. Thus, the data presented in this report present the number of observations of single individuals. It can, however not be excluded that the same individual came back and was filmed and counted again. The data can therefore not be interpreted as true abundances in terms of number of organisms per unit time.

In analysing occurrences of zooplankton, the tedious approach as applied to fish and cephalopods was not feasible. Firstly, often many individuals of the same group appeared in front of the camera, and secondly the duration of stay in front of the camera was very much dependent on the velocity of the surrounding sea water.

Analysis of the SAS-Oden FishCam video recordings

Altogether, the video recordings from the SAS-Oden FishCam mounted on the CTD consisted of 13 CTD profiles down to 1000 m. During the downcast the CTD was stopped every 100 m for 3 minutes down to 800 m and the time to move 100 m downward was ca. 3 minutes as well. The macrozooplankton could only be counted during these three minutes when the CTD had stopped because at other times it moved too fast to recognise any organisms. The field of vision was ca. 100 cm from the camera objective, and red illumination was used to disturb the organisms as little as possible.

The video recordings were analysed manually in a dark room at 0.5× recording speed. The criterion for counting a macrozooplankton individual was that it would be larger than 1 cm on a 27 inch (68.58 cm) computer screen. Close to the camera lens they could be up to ca. 15 cm long. A macrozooplankton individual could move very fast and be captured on the screen for 1-3 seconds, while other individuals moved slower and could remain on the screen for 15-20 seconds (at 0.5× recording speed).

Analysis of the SAS-Oden UVP6 data

Altogether, UVP6 data were collected in 43 CDT casts at 23 stations (maximum sampling depth 567-4 477 m depth) during the SAS-Oden expedition[6]. A total of 71 344 UVP6 images were uploaded to the online annotation tool "Ecotaxa"[75] by Marc Picheral (CNRS-Sorbonne Université, Institut de la Mer de Villefranche sur mer) to establish and ensure the correct processing of the UVP6 images.

EcoTaxa provides random forest algorithms to predict a category for each object, based on optical features of the images[76]. For meaningful predictions, these algorithms are usually based on established learning sets. However, no learning set existed for the UVP6 (the newest UVP model that was used during the SAS-Oden expedition) and those for UVP5 images used during the MOSAiC expedition) did not yield correct predictions. Therefore, the images collected during the SAS-Oden expedition were sorted manually into categories.

Living animals were subdivided into 25 taxonomic categories. Copepods identified on the basis of typical body form and clear antennae were collected in the category "Copepoda". In addition, organisms with antennae and a copepod-like body form but photographed at a suboptimal angle were collected in the category "like<Copepoda" for better predictions

---

[75] https://ecotaxa.obs-vlfr.fr/

[76] Picheral, M. et al. (2022) The Underwater Vision Profiler 6: an imaging sensor of particle size spectra and plankton, for autonomous and cabled platforms. Limnol. Oceanogr. Methods 20:115-129 [https://doi.org/10.1002/lom3.10475]

using the EcoTaxa algorithm. It was verified that these also show copepods, and therefore these two categories were combined for calculations of abundances. In a few images, it was possible to identify the copepod genera *Calanus*, *Paraeuchaeta* and *Metridia*. The number of images for these groups were, however, too low to allow for analysing distribution patterns. In one category all images believed to present organisms, but without being able to identify their taxonomic affiliation, were collected.

Parts of organisms were subdivided into six categories. These categories included body parts such as cnidarian tentacles and other body parts, legs and antennae of crustaceans, and houses of appendicularians. Particles that were embedded in a grey matrix or which were obviously cut out of a larger structure were collected in the category "Appendicularia<over-segmented houses". The name of this category refers to appendicularian houses that were also found complete in the data set (total of 593 houses), with many particles attached. Displaying all images according to their time stamps revealed that the automatic extraction process often produced images presenting these attached particles as single objects while they were in fact part of a larger aggregate.

Non-living particles were subdivided into 7 categories. In the category "Artefacts" grey objects with no discernible features were collected. The largest of these objects were striped hexagons, and it was assumed that most smaller artifacts were subsets of these structures. Moreover, out of the ca. 21 000 images in the "Artefact" category, ca. 20, 000 were found in one single cast (Cast 9 at Station 22). It was therefore safe to remove these images from further analyses as they likely show ice or salt crystals (Marc Picheral, personal communication).

UVP data are commonly extracted via "EcoPart", a tool accessible through the "Ecotaxa" website. EcoPart has been designed (1) to extract raw data of all images taken by the UVP as standardised logarithmic particle size classes (table with abundance and biovolume of objects, combined in size classes from 1 µm to >2.6 mm). As it is not possible to remove any category, the particle data provided by EcoPart always contain artefacts, over-segmented elements and reflections etc., and thus provide a quick but only rough overview of the particle distributions. (2) EcoPart also provides information on the annotated objects, i.e., a table with abundance, biovolume and size of each category. It has to be noted that the UVP stores only images with objects that are larger than ca. 800 µm for later annotation with EcoTaxa; all smaller objects are counted as particles and included in the particle data set but not available for later image analyses. The total number of particles is thus always larger than the number of objects analysed in EcoTaxa.

Following the general procedure for UVP data analysis, two tables thus emerged from the SAS-Oden expedition, one with all raw particle size classes, and a second with data for each category as classified in EcoTaxa. All data were provided as averages L^{-1} in 5-m depth intervals, automatically calculated from the frame rate and the speed of the UVP by EcoPart. These data were then processed as follows: (1) Particle data: Here only size classes larger than 102 µm (equivalent to 3 pixels) were used because smaller pixel sizes reflect noise rather than actual particles (Dr. Andreas Rogge, AWI, personal communication). (2) EcoPart automatically sums up data of categories into taxonomical entities. For example, it provides one averaged value on the abundance of Cnidaria that is calculated from the category "Cnidaria" plus "tentacles<Cnidaria". However, it can be assumed that this severely overestimates the actual values, and the respective averages were corrected, i.e., the number of tentacles<Cnidaria were subtracted. Also, the number of antennae<Crustacea were subtracted from the total abundance of crustaceans, and to derive abundances of Appendicularia, the number of "Appendicularian houses" were subtracted from the overall abundance as the occurrence of houses does not reflect the distribution of appendicularians since dead houses sink quite rapidly into the deeper ocean.

Analysis of the SAS-Oden LOKI data

Altogether, LOKI data were collected in the upper 1000 m of the water column at 11 stations during the SAS-Oden expedition[77]. These 11 LOKI casts resulted in >400 000 images. The first step of image processing was to remove artefacts such as light reflections and objects that were stuck onto the camera lens. Such images were identified and deleted using a Python script. The remaining images were then processed with the instrument's software "LOKI Browser" to extract optical features (e.g., area, grey scale, object length) for each object. Thereafter, the images were converted to Portable Network Graphic (.png) data format and their quality was enhanced by the autocorrection application of the software program "IrfanView"[78]. To detect and delete duplicate images of the same organism or object, the images were processed with the software "ZooMIE"[79]. Finally, 30 168 images were uploaded to and annotated in "Ecotaxa"[80].

To predict categories of the objects on the LOKI images with the random-forest algorithm provided by "EcoTaxa", a learning set for Arctic zooplankton that had been established at the AWI based on previous expeditions was used All predictions were checked and, if necessary, the categories manually corrected. Altogether, 23 310 individual living organisms (including eggs and unknown objects that were likely living organisms) were identified while the remaining images showed dead organisms, bubbles, faecal pellets and other detritus particles.

Since the LOKI uses a plankton net to concentrate zooplankton, the images can be treated as count data, similar to net data, the abundances (number of individuals m^{-3}) were calculated in 10-m depth intervals from the number of organisms identified on images, the net opening ($0.28\ m^2$) and the sampling depth (1000 m), assuming a filtered volume of 280 m^3 at every station. In order to actually measure the filtrated volume, a flow meter to the LOKI net opening was mounted. However, unfortunately this approach did not yield sound data during the SAS-Oden expedition. It also has to be noted that, as expected, the number of organisms in the approximate upper five to ten metres of the water column was high as compared to greater depth. In such cases, the LOKI software often stores full-frame images with many overlapping organisms at a lower frame rate. These full-frame images do not yield meaningful data, and therefore, the total zooplankton numbers at the surface should be interpreted with caution because they are underestimated.

Analyses of hydroacoustic data targeting zooplankton

The basic acoustic data collection methods for zooplankton did not essentially differ from those for fish as described by WP2, except that the data used in the analyses were from other frequencies: ship-mounted EK60/80 MOSAiC, transducers 70, 120, 200 kHz and WBAT mounted on CTD SAS-Oden, transducer 333 kHz.

Zooplankton produce weak echoes and the backscattering strength is highly depended on the frequency used. If a target is too small relative to the acoustic wavelength, then the backscatter will be substantially weak following Rayleigh scattering principle. Therefore, there is a limitation for the usable frequency for the small-sized targets. Another important factor that determines the signal strength is the acoustic properties of the body composition as the backscattering strength is directly related to the density and sound speed contrast of the animal body relative to the seawater. Because most zooplankton do not have large contrast relative to the seawater, their scattering is weak,

---

[77] Snoeijs-Leijonmalm P, et al. (2022) Ecosystem mapping in the Central Arctic Ocean (CAO) during the SAS-Oden expedition. Publications Office of the European Union [https://data.europa.eu/doi/10.2826/958629]

[78] https://www.irfanview.com

[79] Schmid MS, et al. (2015) ZOOMIE v1.0 Zooplankton Multiple Image Exclusion (Version 1.0) Software. [www.zenodo.org]

[80] https://ecotaxa.obs-vlfr.fr/

and this limits the observation range. Therefore, meaningful estimate of the meso-zooplankton from the MOSAiC acoustic data can only be achieved using frequencies 120 kHz and above where the lower frequencies can be used to categorise scatterers into different groups. Due to the weakness of the zooplankton echoes, it is necessary to collect clean signals to achieve the needed signal to noise ratio (SNR). Ideally the background noise should be 10dB lower than the expected strength of the targeted echoes.

Evaluation of the 120 kHz measurements from MOSAiC indicated high noise levels that hindered the use of these recordings as a quantitative index for the zooplankton. Therefore, the focus was on 200 kHz as a quantitative index while 38, 70 and 120 kHz were used to identify potential zooplankton aggregates. 38 kHz was used to identify and eliminate regions where the backscatter is either caused by the sharp density changes such as pycnocline or fish like targets using the condition $\Delta MVBS_{38kHz-200kHz} > 5$. School-like patterns were eliminated. Here, a -68 dB lower threshold was applied for the echogram and using school detection algorithm strong clumps of pixels were eliminated with an expansion mask of 7×3 window (7 vertical samples and 3 pings). The detection limits were 0.5 m per height and 1 m per length. These detections were further visually checked and false positives were manually corrected. For the $\Delta MVBS$ analysis, a consistent rule for 70 kHz and 120 kHz could not be applied but these echograms were used opportunistically to identify/correct layer classification. After careful elimination of non-zooplankton backscatter, the echograms were integrated from the surface to 150 m depth for a distance of 150 m using -88 dB as lower threshold (**Figure 32**). For such threshold, 100 m is the lower limit in most applications. However, the ideal, quiet conditions of RV Polarstern during the MOSAiC enabled the extension to 150 m depth with satisfactory SNR. The export was performed with 1-hour intervals where one estimate (number of individuals per m^2) can be assigned to the corresponding fish backscatter from the 38 kHz data. This metric was then used as an explanatory variable to predict fish abundance for the MOSAiC expedition route.
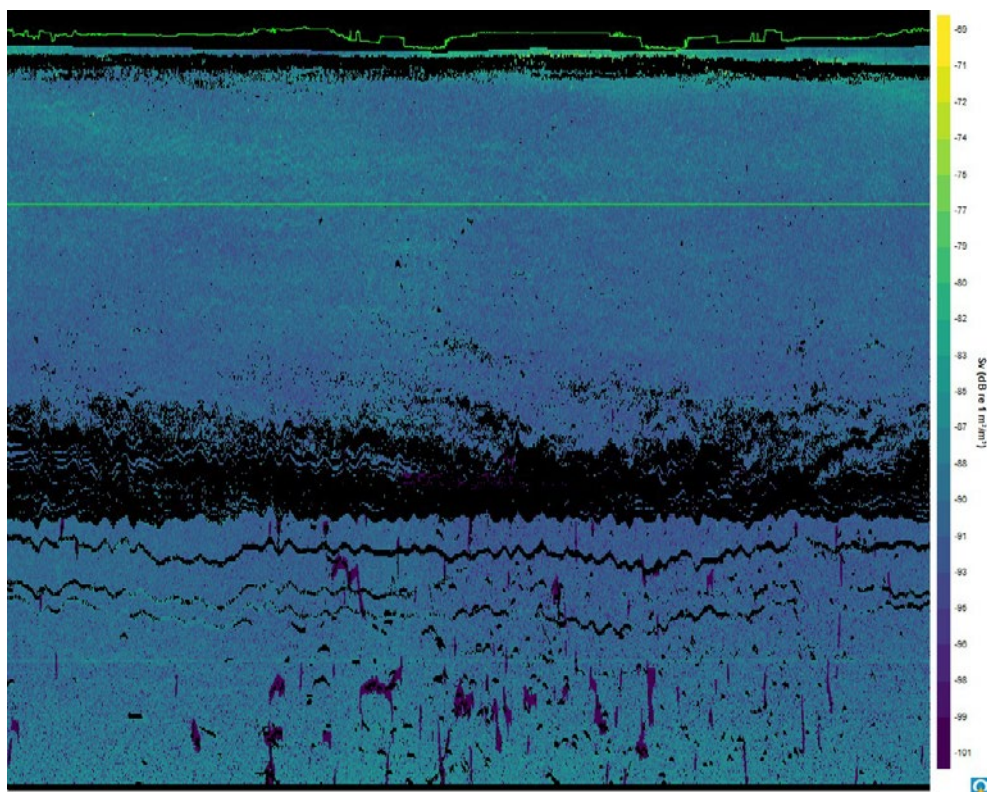
Figure 32: The first 150-m section of the 200 kHz transducer after cleaning for ΔMVBS38kHz-200kHz >5 and removal of fish-school-like small patches.

Acoustic measurements on carried out using a stand-alone echosounder (WBAT-Kongsberg Maritime AS) equipped with a 333 kHz transducer (ES333-7CDK split beam) operating in wideband mode (320–420 kHz, 2.028 pulse duration (ms) and 2 Hz pulse). The transducer was mounted on the bottom of the CTD, facing sideway. At each station the WBAT was lowered to the target depths (maximum 1000 m) and continuous profiles were recorded in combination with CTD measurements. WBAT calibration was performed with a 22.0 mm tungsten carbide (6% cobalt bond) calibration sphere[81]. Broadband calibration parameters were collected using the EK80 Calibration Wizard of the Simrad EK Mission Planner version 3.3.0 and processed in Echoview (version 12.1, Echoview Software Pty Ltd, Hobart, Australia, www.echoview. com).

The maximum recording range was set at a range of 50 m considering the limitation of the signal-to-noise ratio (using 10 dB as threshold). During analysis, this range was further reduced to 38 m to reduce overlaps between the vertical samples due to beam expansion. In addition, the first 2 m of the transducer were also removed due to potential evasion effects. Actual depth was measured by the CTD pressure sensor associated with WBAT measurements based on ping time with millisecond precision. To assess vertical size distribution, individual targets were tracked from single target echograms using a similar method to fish tracking, as described in WP2. Here, care was taken to remove the strongest targets that are potentially fish or cannot be classified as zooplankton. For this, initial manual clean-up was performed for visually identifiable noise as described in WP2, including noise patterns such as bubble/ice entrainment near surface, backscatter from surface/bottom, visually identifiable noise spikes. Also, all upcast recordings were removed due to possible disturbance effects.

After the manual cleaning of the data an upper threshold of -55 dB was applied to the TS echogram. Then these masked sections were further magnified using a 21×5 (21 vertical samples and 5 pulses). After removal of such strong regions, individual targets were detected using Echoview's target detection algorithm and again filtered by a secondary threshold of -62 dB. From this clean, filtered echogram, individual tracks were detected using Echoview's fish tracking algorithm.

While the target tracking is useful for estimating the size distribution of the targets and acquiring a rough number calculation, the echo integration provides a more accurate bulk estimate for the biomass of the whole assemblage. For echo integration a similar cleaning procedure as in the target tracking was applied, where the window size was 25×9 for strong regions above the threshold of -62 dB, and a secondary threshold of -68 dB was applied to further remove strong echoes. The reason for using this conservative approach is that it is more susceptible to the inclusion of non-targeted samples, whereas target tracking excludes most such samples due to selective criteria during target tracking.

One way of estimating zooplankton abundance is to divide the mean volume backscattering strength in a unit water volume by the mean target strength of the individual targets that are responsible for the echoes. This method requires, individual animals to be separately detected using advanced methods such as target tracking. Once the underlying TS distribution of the individuals is captured, then the mean TS can be used to convert volume backscatter to the number of individuals. Such a method would require close-range measurements with high acoustic frequencies, such as the 333 kHz profiles collected in this study. Another method is to acquire a representative biological sample from the sampled volume using a net and estimate the mean TS of the assemblage in that unit volume using physic-based TS models taking the size distribution and acoustic characteristics of the taxonomic groups. A third method assesses the relationship between the volume backscattering strength and the biomass estimate from the net samples and uses the volume backscattering measurement as a scaling factor. In this chapter results based on the first and the third method are provided.

---

[81] Demer DA, et al. (2015) Calibration of acoustic instruments. ICES Cooperative Research Report 326:1-133 [https://ices-library.figshare.com/articles/report/Calibration_of_acoustic_instruments/19056617]

The acoustic area scattering coefficients were extracted per depth layer and time interval using echo integration. The results were combined and interpreted together with the WP3 LOKI, UPV, and FishCam data and the WP4 community composition and biomass data based on the multinet measurements.

3.5. Results and discussion of WP3

Fish and squid along the MOSAiC expedition route (MOSAiC FishCam)

The images extracted from the video footage often showed blurry structures, but it was also possible to identify fish and squid (**Figure 33 A, D**). Even smaller organisms such as copepods, chaetognaths, ctenophores and siphonophores (**Figure 33 B, C, E, F**) were present on the extracted images. A major obstacle was that the FishCam did not allow to determine organism sizes which made the identification of taxa, especially in zooplankton very difficult, if not impossible, on most of the images. Copepods, for example, appeared to be very large if moving right in front of the camera lenses while jellyfish appeared small if they were far away from the FishCam.

Analysing all images extracted from the video footage via algorithms yielded only six observations of fish individuals (**Table 25**). The first image showed fins, and the video revealed that this fish was only visible at the edge of the video frame. Thus, this individual could not be further identified. Three myctophid individuals were observed on 16 and 18 December 2019 and on 23 January. On 28 January, *Gadus morhua* was found twice, about half an hour apart, at 3:26 am and at 3:58 am, and it is possible that here the same individual passed the camera twice. Throughout February and early March, there were no more sightings, and after March 11, the camera could not be deployed anymore due to unfavourable ice conditions. All fish were found in the north-eastern part of the Amundsen basin (**Figure 34**). Altogether, the number of fish observations is far too low for drawing conclusions about fish distribution patterns in the CAO.
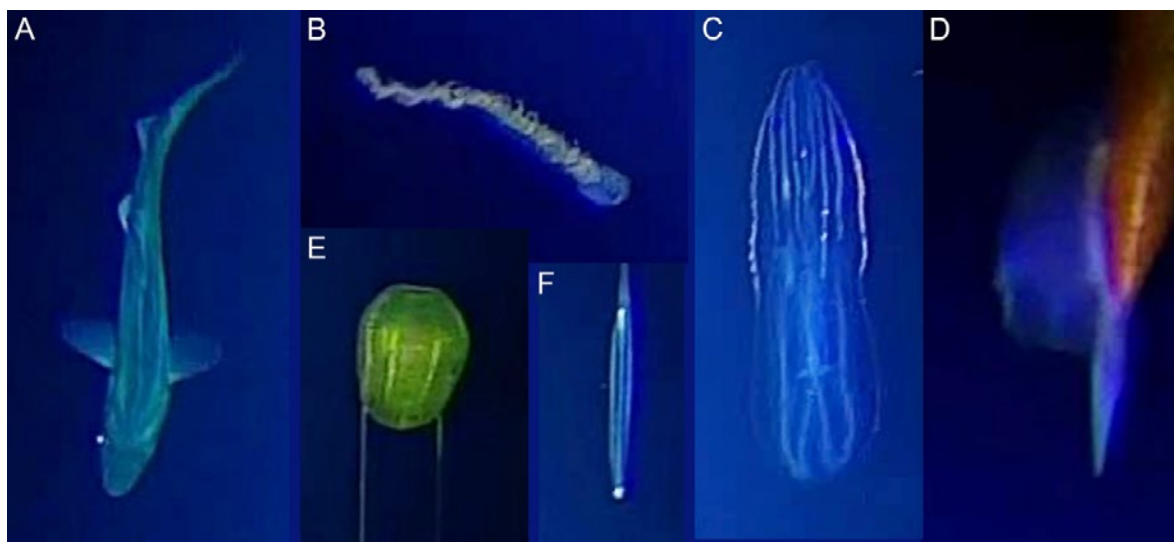
Figure 33: Selected images extracted from the MOSAiC FishCam videos by an automatised video analysis script (A) *Gadus morhua***, (B) Siphonopora, (C) Ctenophora. (D) squid. (E) Ctenophora. (F) Chaetognatha. Please note that, due to the algorithms applied, the size of the images does not reflect the size of the organisms, and no scales can be included.**

Table 24: Fish observations during the MOSAiC expedition using a deep sea camera system (FishCam, manufactured by MacArtney, Kiel, Germany). The data present observations of single fish individuals, including date, time and location, as well as the taxon identified from the images, if possible.

Date	Time	Depth	Latitude (°N)	Longitude (°E)	Species
15.12.19	20:09:41	369-376 m	86.62	117.64	not identifiable
16.12.19	19:45:26	369-376 m	86.60	115.67	*Benthosema glaciale*
18.12.19	15:55:36	369-376 m	86.68	113.42	*Benthosema glaciale*
23.01.20	05:17:13	369-376 m	87.45	94.18	*Benthosema glaciale*
28.01.20	03:26:35	213-215 m	87.45	95.85	*Gadus morhua*
28.01.20	03:58:25	213-215 m	87.45	95.85	*Gadus morhua*
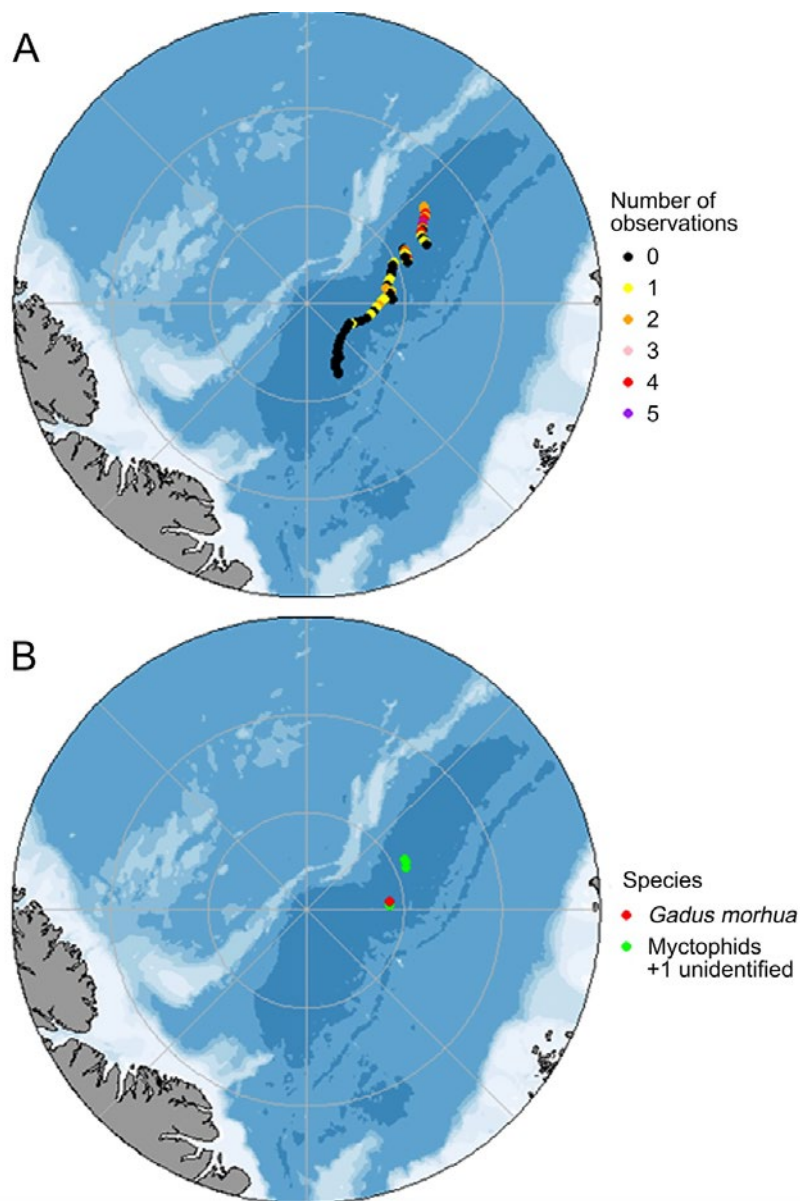
Figure 34: Observations of nekton along the MOSAiC expedition route by automatic analysis of video recordings taken by a deep sea camera system (FishCam, manufactured by MacArtney, Kiel, Germany). (A) Number of observations of armhook squid (*Gonatus fabricii*). (B) Two observations of Atlantic cod (*Gadus morhua*) on one day (red dot) and three observations of myctophids and one unidentified fish (green dots).

In the already published part of the study in which 9.3% of the video recordings were included[82], four fish and 39 individuals of *Gonatus fabricii* (armhook squid) were identified by manually analysing 180 video recording hours (**Table 25**). A comparison between the two methods showed that the automated method is less effective than manual counts, but does capture the general pattern of squid occurrence with more squids being observed during the first months of the MOSAiC expedition. Also, the real-time observation of the videos yielded more fish than the automatic extraction of fish images. This may be explained by the overall low numbers of fish individuals compared to the background noise caused by zooplankton. If video monitoring was to be included in monitoring programs, it is suggested to implement automatic processing of the video despite its shortcomings since the time for the analysis was only about one-tenth of that of analysing the videos manually (in this case three months instead of 30 months).

Table 25: Comparison between manual and automatic recording of nekton for the six five-day periods that were manually analysed.

Time period	Minutes manual	Minutes automatic	Fish manual	Fish automatic	Squid manual	Squid automatic
24-28.10.2019	2400	6197	0	0	10	13
02-06.11.2019	2400	7200	0	0	12	7
13-17.12.2020	2400	9041	2	2	10	11
17-21.01.2020	1200	10850	0	0	2	6
05-09.02.2020	1200	7690	1	0	0	3
06-10.03.2020	1200	7304	1	0	0	0
	10800	48282	4	2	34	40

Armhook squid (*Gonatus fabricii*) was encountered 76 times from the end of October 2019, when the camera was deployed, until 11 March, when it was recovered (**Figure 35**). Unfortunately, the camera was not operative from 12 November to 12 December but when it was redeployed from 12 to 21 December, squids were found on the videos again on an almost daily basis. Thereafter, from 5 January to 11 March, the system worked continuously except for four consecutive days (30 January – 2 February 2020) and one single day (10 February). During this period, the frequency of armhook squid sightings decreased in terms of both days with positive identifications and number of sightings per day was low. The last day on which a squid appeared on video was 21 February 2020. With regard to the spatial distribution, it appears that armhook squid was relatively frequent in the eastern part of the Amundsen Basin, while their numbers decreased towards the middle and western part of the Amundsen Basin (**Figure 34**).

It has to be noted that all data above present total numbers of observations per day but do not account for the number of minutes that the two cameras of the FishCam were filming on the respective days. As the light(filming):dark(non-filming) periods were alternated during the expedition, this number varied between 139 min (only on day 1) and 2640 min. Projecting the number of observations on a 24-hour period (**Table 26**) yielded extremely high values on the first two deployment days (23 October 2019: 20.7 squid observations per day, 24 October 2019: 12 squid observations per day). Such calculations can overestimate the occurrence of armhook squids when the video recording time was short while long video footage may underestimate their occurrence. Moreover, as with the Atlantic cod, it cannot be excluded that the same individual passed the camera twice or even more often. Thus, these data can provide only a rough estimate of the distribution of armhook squid in the CAO.

---

[82] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]
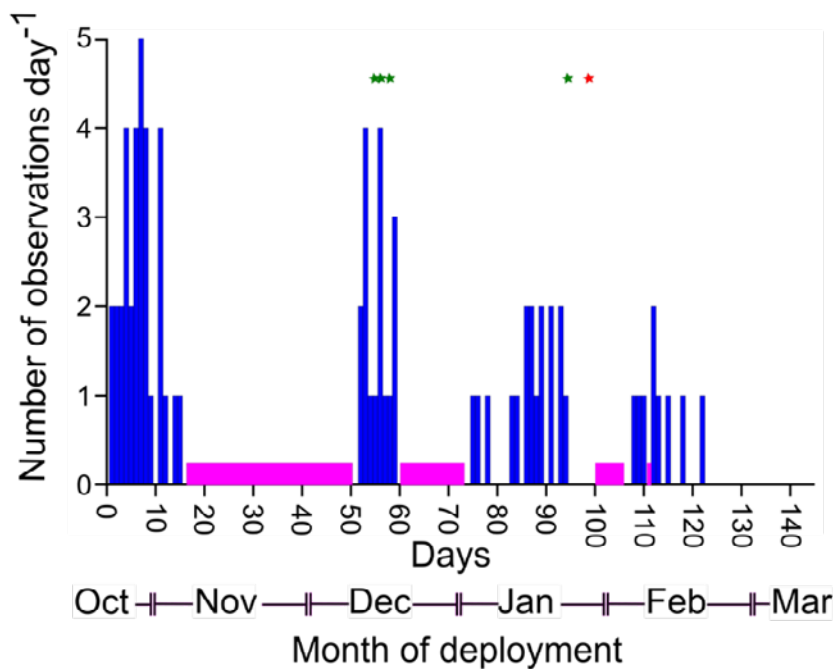
Figure 35: Observations of nekton during the MOSAiC expedition from 23 October to 11 March analysed by automatic extraction of images from videos taken by a deep sea camera system (FishCam, manufactured by MacArtney, Kiel, Germany). Bars present the number of sightings of armhook squid within 24 hrs; green stars indicate an observation of one fish day^{-1} (unidentified or myctophid); red star represents two observations of Atlantic cod *Gadus morhua* at one day. When the camera did not deliver data due to unfavourable conditions or technical problems, days are marked in pink.

Particle distribution along the MOSAiC expedition route (UVP6)

The size of most particles (39%) recorded with the UVP6 on the CTD during the SAS-Oden expedition ranged between 102 and 128 µm (**Figure 36**), followed by the size-fractions 128-161 µm (28%), and 161-203 µm (19%). The next largest size classes contributed considerably less, with 7% (203-256 µm), 4% (256-323 µm), 2% (323-406 µm) and 406-512 µm. Particles larger than that µm were rare (in total 0.5%).

The number of particles was relatively high in the surface layer (0-20 m), with the highest values found at Stations 22, 44, 47, 50 and 57 (**Figure 37**). However, the particle counts at the surface are likely too high as here for example bubbles might have been counted but cannot be omitted from the data set (see Methods). With certainty, the high abundance at Station 22 can be explained by the occurrence of salt or ice crystals. With increasing depth, the concentrations decreased sharply and reached minimum values as low as <5 particles L^{-1} between ca. 100 and 200 m. At most stations, a second peak in particle abundance was found between 200 and 500, which could reflect higher concentrations in the Atlantic water mass. An exception was Station 50 on the Greenland Shelf Slope, where Polar water prevailed. At Stations 20, 22, 24, and 25, which were located in the Amundsen Basin but relatively close to the Lomonosov Ridge, the particle concentrations increased at 1500-2000 m depth, and remained comparably high to depths of 3000 m and more. At the other stations, the number of particles was low in greater depth, and only increased close to bottom, possibly due to the occurrence of resuspended material.

Table 26: Observations of armhook squid (*Gonatus fabricii*) during the MOSAiC expedition using a deep-sea camera system (FishCam, MacArtney, Kiel, Germany). Presented are date and daily means of latitude and longitude, day of study, number of observations on images extracted from recordings of two cameras (minutes per video day) as well as number of squids per 24 hours calculated from the number of observations and the duration of the recording.

Date	Latitude (°N)	Longitude (°E)	Day	Number of observations	Minutes per video day	Number of squids per 24 hours
20191023	85.349	129.884	1	2	139	20.72
20191024	85.406	128.893	2	2	240	12.00
20191025	85.446	127.894	3	2	1127	2.56
20191026	85.458	127.387	4	4	2660	2.17
20191027	85.500	126.858	5	2	1130	2.55
20191028	85.587	126.140	6	4	1040	5.54
20191029	85.666	125.210	7	5	1440	5.00
20191030	85.726	124.266	8	4	1440	4.00
20191031	85.775	123.429	9	1	1440	1.00
20191102	85.859	122.401	11	4	1440	4.00
20191103	85.905	121.273	12	1	1440	1.00
20191105	85.939	118.469	14	1	1440	1.00
20191106	85.919	117.456	15	1	1440	1.00
20191213	86.596	119.147	52	2	970	2.97
20191214	86.612	118.453	53	4	1155	4.99
20191015	86.618	117.913	54	1	1920	0.75
20190816	86.604	116.414	55	1	2640	0.55
20190617	86.637	114.563	56	4	2356	2.44
20190418	86.679	113.543	57	1	2640	0.55
20191219	86.679	112.953	58	1	2640	0.55
20191220	86.681	112.581	59	3	2640	1.64
20200105	87.069	115.321	75	1	1710	0.84
20200106	87.103	115.161	76	1	2200	0.65
20200108	87.116	115.126	78	1	2640	0.55
20200113	87.337	107.081	83	1	2640	0.55
20200114	87.410	104.999	84	1	2640	0.55
20200116	87.556	101.474	86	2	660	4.36
20200117	87.500	99.259	87	2	2115	1.36
20200118	87.407	98.341	88	1	1712	0.84
20200119	87.402	98.299	89	2	2220	1.30
20200121	87.487	96.100	91	2	2420	1.19
20200123	87.441	93.933	93	2	2190	1.32
20200124	87.419	93.108	94	1	1780	0.81
20200207	87.642	93.799	108	1	1790	0.80
20200208	87.679	93.745	109	1	1750	0.82
20200209	87.733	92.910	110	1	720	2.00
20200211	87.825	89.802	112	2	1980	1.45
20200212	87.883	87.420	113	1	1616	0.89
20200214	88.030	80.511	115	1	1928	0.75
20200217	88.074	78.831	118	1	1643	0.88
20200221	88.503	66.484	122	1	540	2.67
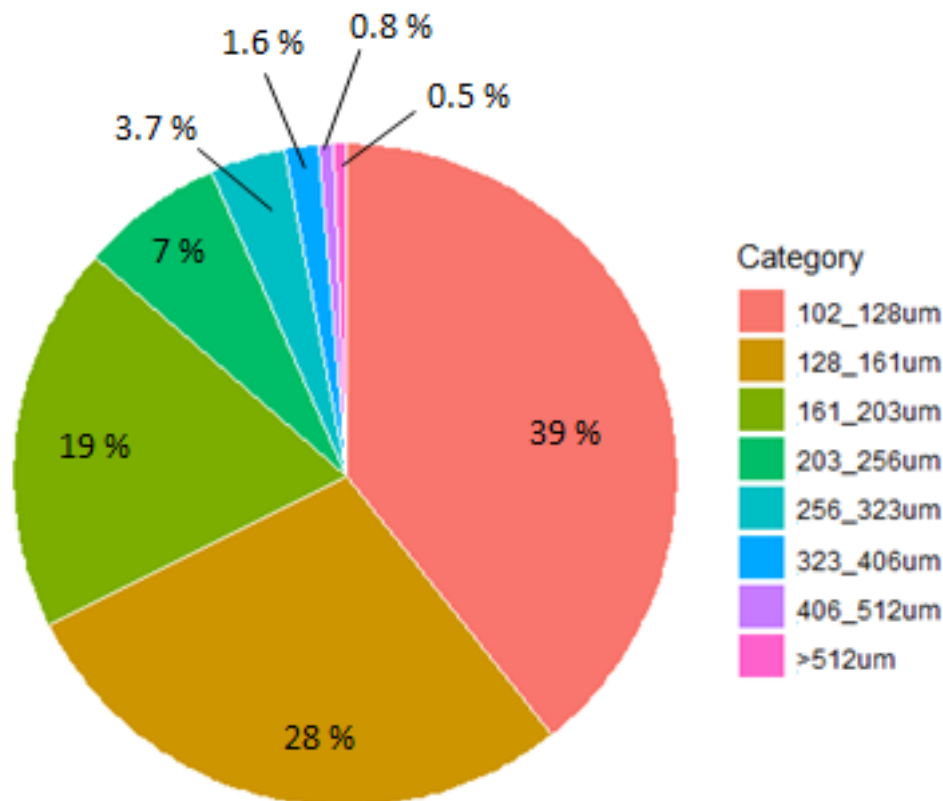
Figure 36: Particle size class composition derived from UVP6 casts during the SAS-Oden 2021 expedition.

Zooplankton community composition along the MOSAiC expedition route (UVP6, LOKI)

The UVP6 recorded images over the entire depth range of a sampling station, i.e., from the surface to the bottom (maximum depth ca. 4500 m), while usage of the LOKI was restricted to the upper 1000 m of the water column due to the technical properties of the system. To be able to compare the community composition obtained from each of the optical systems, relative abundances from the entire cast and from data between 0 and 1000 m of depth were calculated for the UVP6 data. Altogether, the 1000-m casts comprised on average 1936 organisms per station for the UVP6 (23 stations) and 1897 organisms per station for the LOKI (11 stations).

Copepods were the most abundant taxonomic group in both the UVP6 and LOKI measurements (**Figure 38**), contributing with 71.5% (UVP6) and 80% (LOKI) to community composition at 0-1000 m of depth. Rhizaria were the second-largest group on LOKI images (9.7%), but not on the UVP6 images (0-1000 m: 5.5%). Instead, Cnidaria were more frequent on UVP6 images (0-1000 m: 9.9%) than on LOKI images (1.5%). Chaetognaths and ostracods were found with similar percentages on UVP6 and LOKI images, 1.6-1.7% and 2.8-3.9%, respectively. Differences between UVP6 0-1000 m and 0 m - bottom were not very large, except that the% unknown was larger for the entire profile (7.2%) than for the 0-1000 m section (3.7%). Due to the better resolution of the images, LOKI allowed for 98% of the organisms to be assigned to a taxonomic category (1.7% unknown) while on the UVP6 images, 3.7% could not be identified (0-1000 m comparison).
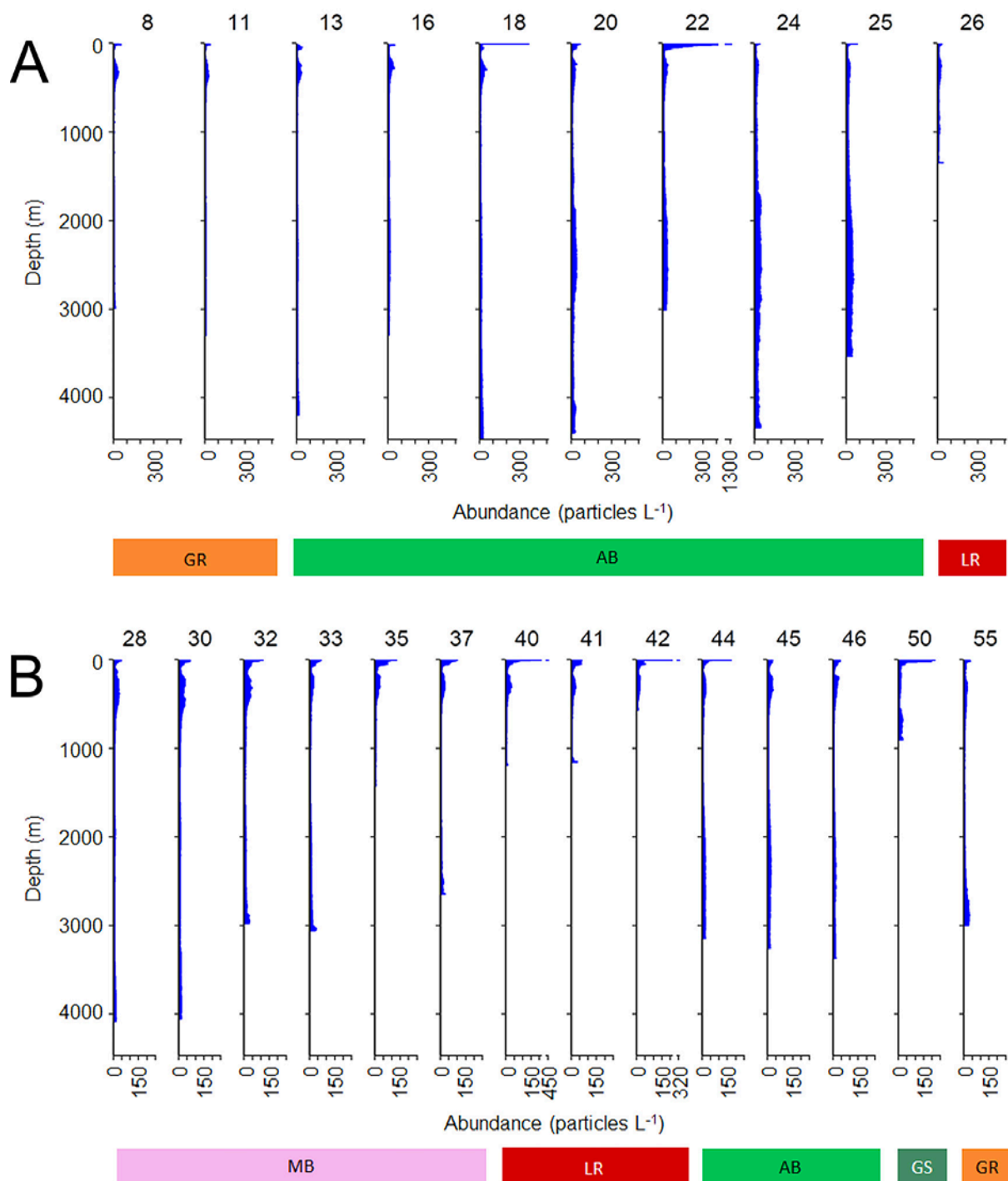
Figure 37: Vertical particle distribution (including non-living and living objects) derived from UVP6 casts from the surface to the bottom during the SAS-Oden 2021 expedition. Presented are numbers of particles L^{-1} averaged over 5-m depth intervals. (A) Northwest transect (1) from Gakkel to Lomonosov Ridge, (B) Southeast transect (2) from Lomonosov to Gakkel Ridge. Profiles are indicated by station number. GR = Gakkel Ridge, AB = Amundsen Basin, LR = Lomonosov Ridge, MB = Makarov Basin, GS = Greenland Shelf. Note the different scales on the X-axes.

The high resolution of the LOKI images allowed to further separate the copepods into lower taxa showing that within this group, Calanoida were dominant (76%), followed by Cyclopoida (23%). Harpacticoida and Mormonilloida were rare (<1%). Within the Calanoida group, large and/or abundant individuals could often be identified to genus or species level, e.g., *Calanus* spp., *Metridia longa*, *Paraeuchaeta* spp., *Scaphocalanus magnus*, *Scaphocalanus* sp., *Microcalanus* spp., and *Spinocalanus magnus*. Smaller and/or rare species were grouped as "Calanoida". The Cyclopoida consisted of 90% Oncaeidae and 10% *Oithona similis*.
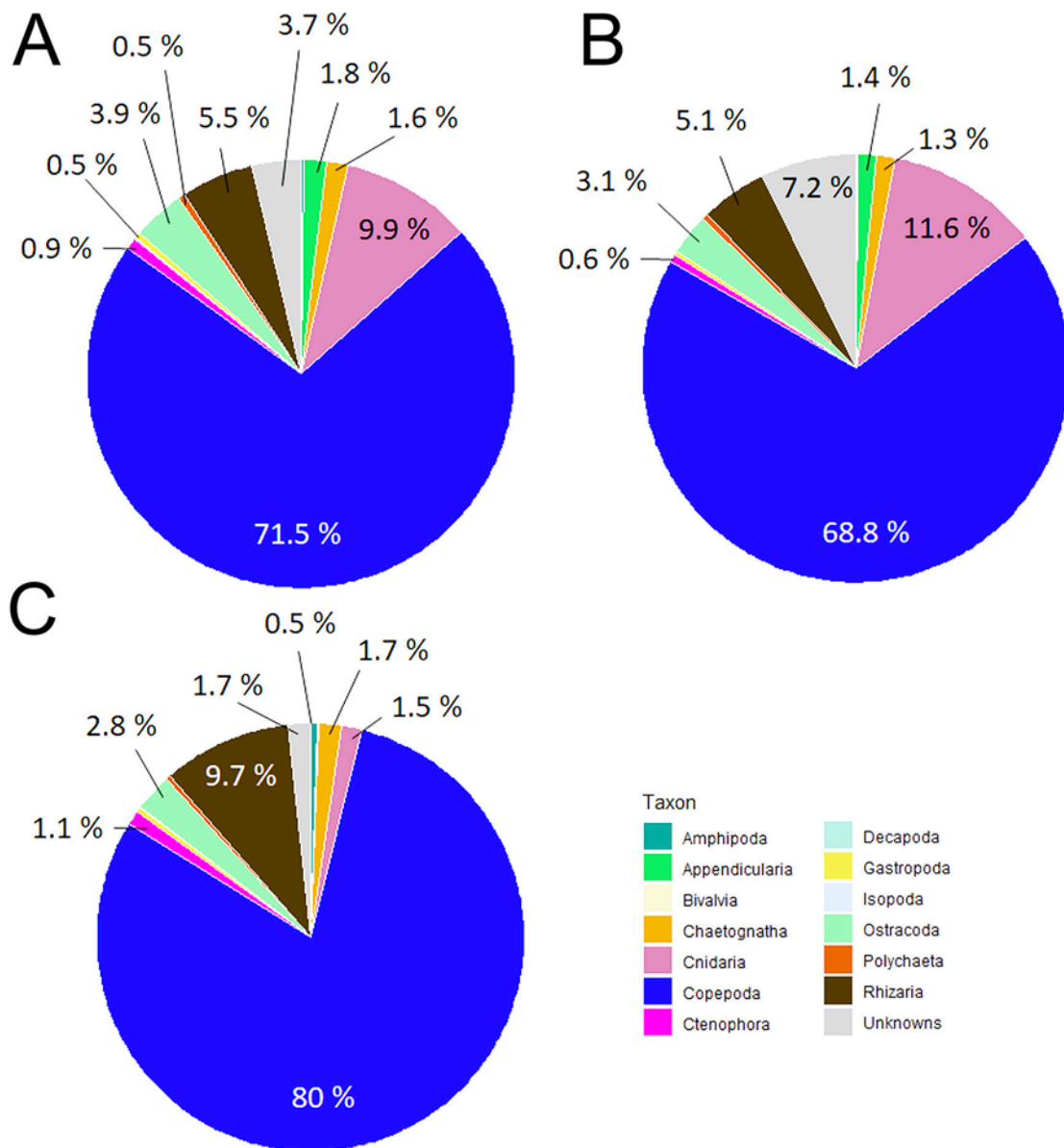
Figure 38: Mesozooplankton community composition measured by image analysis. The images were taken by the optical instruments UVP6 (23 stations) and LOKI (11 stations) during the SAS-Oden 2021 expedition. The percentages represent the relative abundances of higher taxa that contribute ≥ 0.5% to the community composition. Unknowns include living organisms that could not be further identified. (A) UVP6, 0-1000 m of depth. (B) UVP6, 0 m - bottom (C) LOKI, 0-1000 m of depth.

Zooplankton vertical distribution

Since the UVP6 was mounted on the ship CTD and thus frequently deployed (26 stations to full depth), it provided the most complete spatial overview of the zooplankton vertical distribution, albeit with taxonomic limitations. In general, abundances were comparably high during the first transect (Gakkel Ridge to Lomonosov Ridge), and considerably lower in the Makarov Basin and on the Greenland Shelf. The zooplankton depth distribution followed a similar pattern at all stations, with usually relatively high abundances from the surface down to ca. 50 m (maximum of 451 individuals m^{-3} at Station 18) and lower abundances at greater depth (**Figure 39** and **Figure 40**). However, in the Makarov Basin differences between the surface and deeper water did not exceed a factor two, much less pronounced than the differences at Stations 8 and 13 on the Gakkel Ridge and Stations 16, 18, 22 and 24 in the Amundsen Basin. At the Gakkel Ridge and in the Amundsen Basin, organisms were also relatively frequent in most of the 5-m intervals between ca. 200 and 500 m. Many of these 5-m intervals were devoid of zooplankton in the Makarov Basin, and if organisms were present they had lower abundances than in the Amundsen Basin.

With regard to the community composition in the upper 1000 m analysed from the UVP6 data, copepods dominated most depth intervals, especially in areas, where abundances were relatively high (**Figure 40**). During Northwest Transect 1 from the Gakkel Ridge to the Lomonosov Ridge, appendicularians were often found in the surface layer, especially close at the Gakkel Ridge, while ostracods resided deeper, at 50-250 m of depth. Chaetognaths were also mostly found in deeper water, which is probably related to their predatory feeding strategy. Copepods were less important during Southeast Transect 2 from the Lomonosov Ridge to the Gakkel Ridge. Here, Ctenophora and Cnidaria contributed relatively more to the mesozooplankton community composition at greater depth, especially at Station 42 on the Lomonosov Ridge.

Similar to the UVP6, the depth distribution of the mesozooplankton organisms analysed from the LOKI data generally showed the highest (yet underestimated) abundances at the surface. Smaller peaks in abundance occurred between ca. 150 and 500 m of depth, especially at Stations 22, 26, 33, 38 and 57, whereas zooplankton abundances below 500 m were generally very low (**Figure 41**). With regard to depth distribution, less distinct patterns emerged with the LOKI than with the UVP6 since LOKI captures copepods well but not jellyfish. Consequently, in most depth intervals, copepods dominated the communities, often contributing with >90% to the total mesozooplankton abundance (**Figure 41**). As with the UVP6, ostracods increased in importance in deeper water. Cnidaria, a group that was well represented on UVP6 images, only provided considerable proportions at ca. 700-750 m depth at Station 53 on the Morris Jesup Rise. With regard to abundance, the LOKI detected more organisms in the Makarov Basin than in the Amundsen Basin, but this is only based on data from one station (Station 22) and this dominance may thus not be representative for the area.

The LOKI allowed for analysing the copepod community in greater detail that the UVP6 (**Figure 42**), and detected more pronounced regional differences in the potential mesozooplankton prey field for fish. The data indicate that most of the large, lipid-rich *Calanus* species resided in the upper 100 m of the water column at almost every station, except for Station 22 in the Amundsen Basin and Station 57 on the Yermak Plateau, but they also contributed to the copepod populations in deeper water. Another important copepod genus observed was *Metridia*, an omnivorous copepod that was mainly found below the upper surface layer, probably inhabiting the Atlantic Water Layer. Below that, *Scaphocalanus* dominated on the Gakkel Ridge, in the Amundsen Basin and on the Lomonosov Ridge, followed by genera of the Aetideidae, while *Spinocalanus* was more abundant in the Makarov Basin. Oncaea, a small cyclopoid copepod (<1 mm) that is frequently found in deeper water layers, e.g., in Fram Strait (personal observation), contributed considerably to the population in the upper 200 m on both the Morris Jesup Rise (Station 53) and the Yermak Plateau (Station 57).
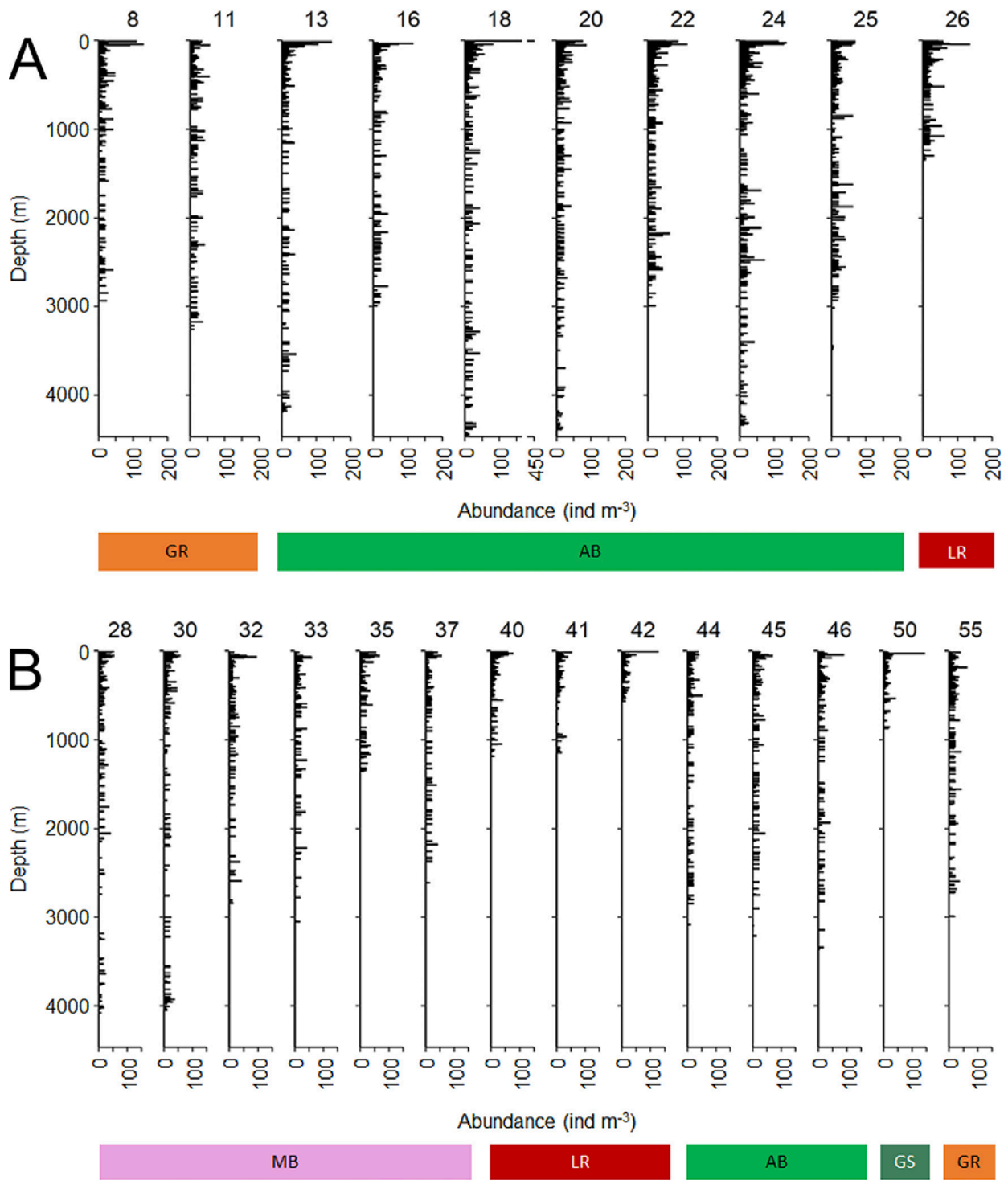
Figure 39: Mesozooplankton distribution from UVP6 images collected during the SAS-Oden 2021 expedition for complete depth profiles. Presented are total abundance (mean number of individuals m^{-3} averaged over 10-m intervals). (A) Northwest Transect 1 from the Gakkel Ridge to the Lomonosov Ridge. (B) Southeast Transect 2 from the Lomonosov to the Gakkel Ridge. Profiles are indicated by station number. GR = Gakkel Ridge, AB = Amundsen Basin, LR = Lomonosov Ridge, MB = Makarov Basin, GS = Greenland Shelf. Note the different scales on the X-axes.
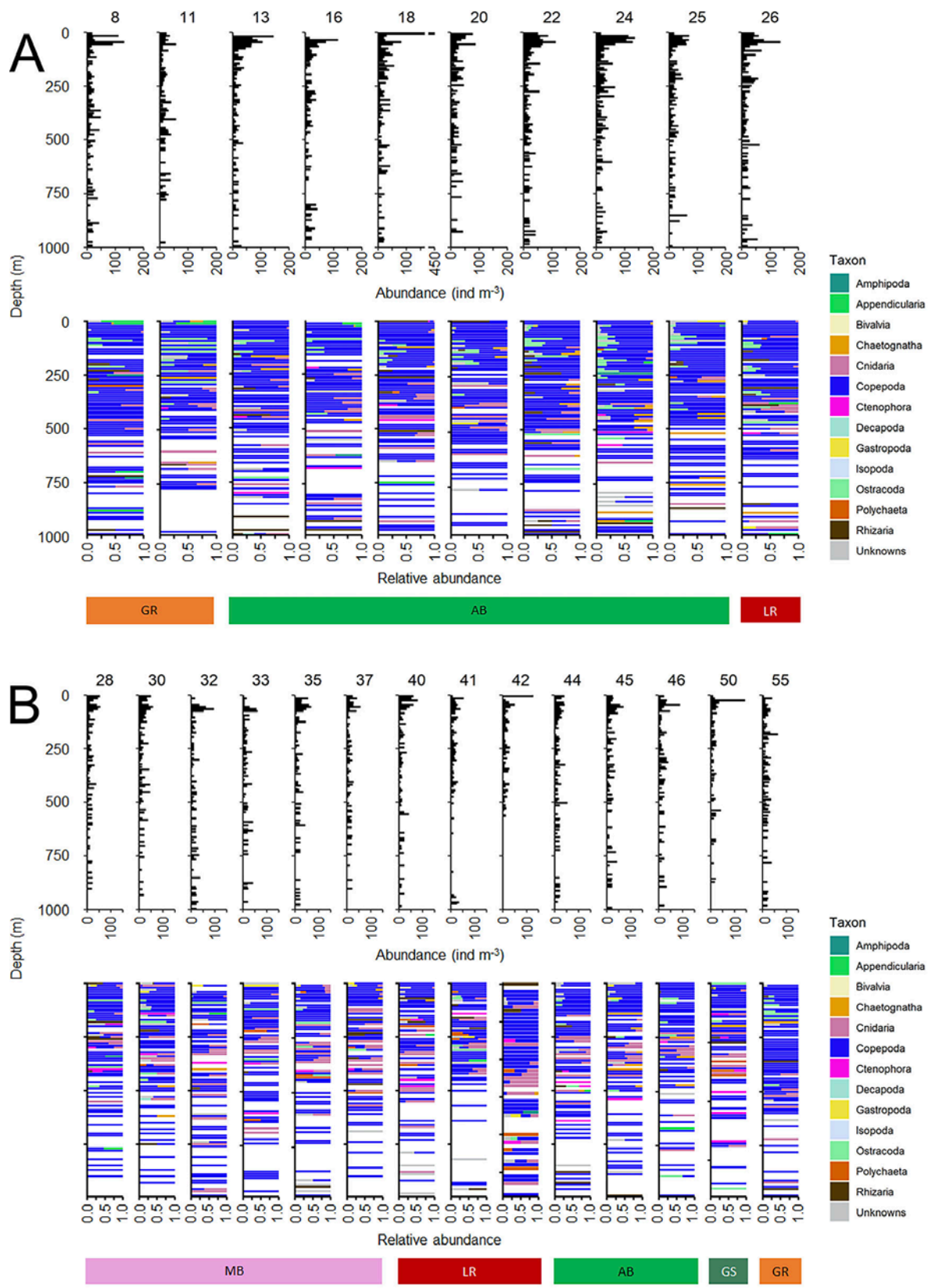
Figure 40: Mesozooplankton distribution and community composition from UVP6 images collected during the SAS-Oden 2021 expedition for the 0-1000 depth interval. Presented are total abundance (mean number of individuals m⁻³ averaged over 10-m intervals) and relative abundance of each taxon. (A) Northwest Transect 1 from the Gakkel Ridge to the Lomonosov Ridge. (B) Southeast Transect 2 from the Lomonosov to the Gakkel Ridge. Profiles are indicated by station number. GR = Gakkel Ridge, AB = Amundsen Basin, LR = Lomonosov Ridge, MB = Makarov Basin, GS = Greenland Shelf. Note the different scales on the X-axes.
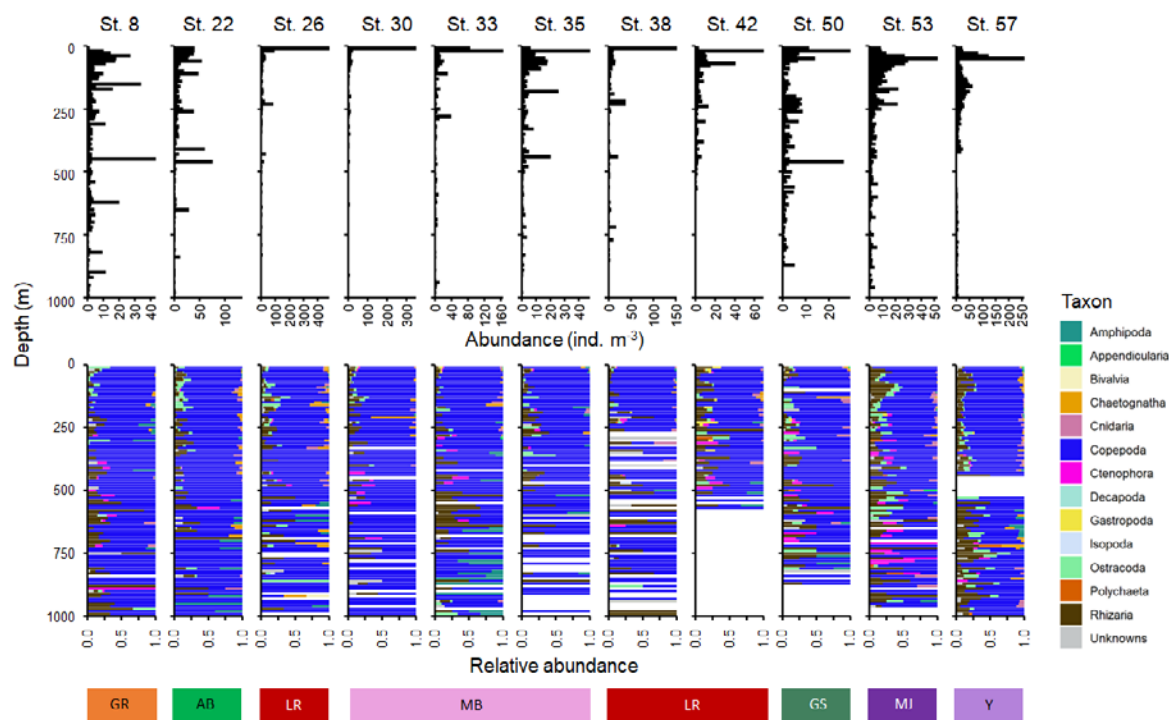
Figure 41: Mesozooplankton distribution and community composition from LOKI images collected during the SAS-Oden 2021 expedition. Presented are total abundance (mean number of individuals m⁻³ averaged over 10-m intervals) and relative abundance of each taxon. Profiles are indicated by station number. GR = Gakkel Ridge, AB = Amundsen Basin, LR = Lomonosov Ridge, MB = Makarov Basin, GS = Greenland Shelf, MJ = Morris Jesup Rise, Y = Yermak Plateau. Note the different scales on the X-axes.
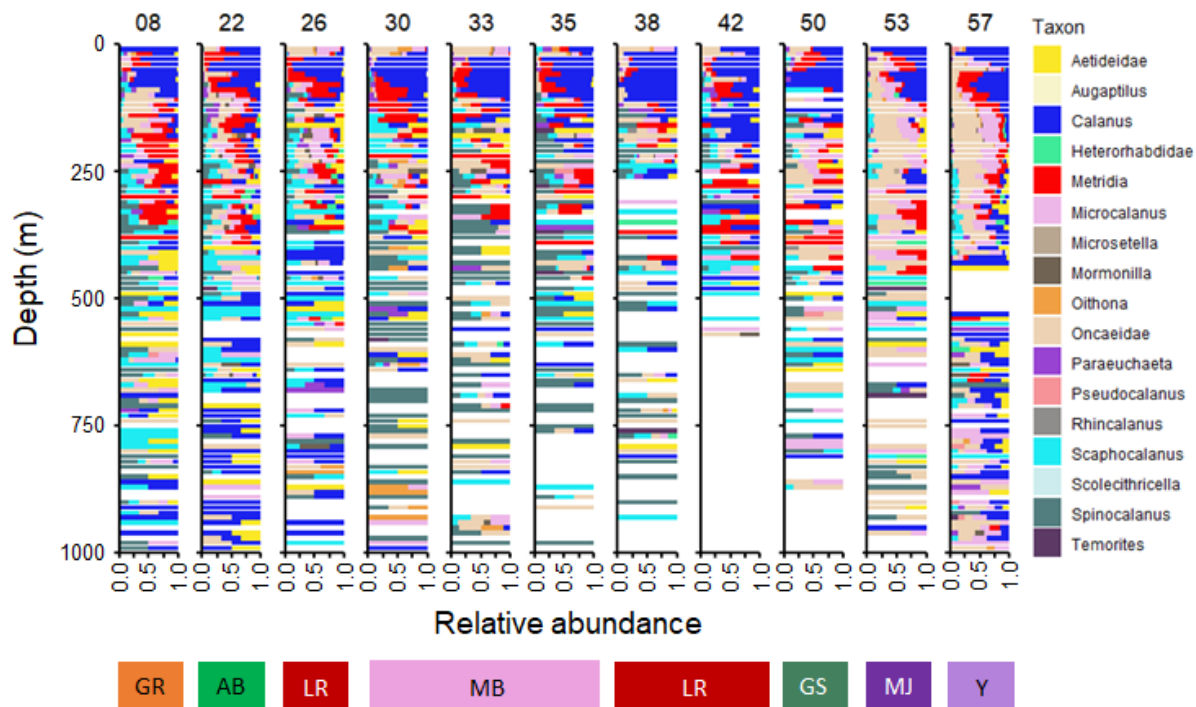
Figure 42: Mesozooplankton distribution and community composition from LOKI images collected during the SAS-Oden 2021 expedition. Presented are total abundances (mean number of individuals m⁻³ averaged over 10-m intervals) and relative contributions of each taxon. Profiles are indicated by station number. GR = Gakkel Ridge, AB = Amundsen Basin, LR = Lomonosov Ridge, MB = Makarov Basin, GS = Greenland Shelf, MJ = Morris Jesup Rise, Y = Yermak Plateau. Note the different scales on the X-axes.

84

Depth distribution of macrozooplankton recorded by the FishCam on the CTD (SAS-Oden)

The 13 depth profiles recorded on video by the FishCam with red light mounted on the CTD during the SAS-Oden expedition showed that the macrozooplankton generally occurred from the surface down to at least 800 m of depth (**Table 27**). Often the maximum abundance of the macrozooplankton was at 100 m and in more than half of the profiles there was also a clear maximum at the lower end of the DSL (600-700 m). No fish was observed on the FishCam on the CTD during the SAS-Oden expedition.

The lower maximum was caused by amphipods (and perhaps some decapods) that were attracted by the light at 500-800 m of depth, even if red light was used with the intention to disturb the marine fauna as little as possible. At 100-400 m no sign of light attraction was observed at any of the stations, suggesting that different species occur in the surface layer and the deep DSL. At the stations with a lower maximum, numerous amphipods turned up at 500-800 m of depth about ca. 2.5 minutes after the CTD had stopped. Probably these were amphipods from the outer sphere around the light where the light was still strong enough to cause attraction. The amphipods came close to the camera objective, almost attacking it (**Figure 43**). A lower maximum was absent in two regions where the Atlantic Water Layer was nearly absent: (1) Station 26 (Lomonosov Ridge) and Station 30 (Makarov Basin), and (2) Stations 50 and 53 (Greenland shelf) (**Table 27**).

At 100-400 m amphipods occurred as well, but they were not as dominant as at 500-800 m of depth and did not seem to be attracted to the light. At shallower depths also some large copepods (*Calanus hyperboreus*), jellyfish (hydrozoa, ctenophores) and some chaetognaths and ctenophores were observed, i.e., a similar community composition as observed on the MOSAiC FishCam video recordings[83]. Generally, the animals were less identifiable on the SAS-Oden FishCam than they were on the MOSAiC FishCam (different camera systems were used), and only the total number of individuals was recorded, i.e., no taxonomic distinctions were made because that would have given unreliable results.

The discrepancy between **Table 27** (deep maximum of amphipods) and **Figure 47** (no deep maximum of zooplankton) is most probably caused by the light attraction. However, deeper zooplankton maxima were detected by hydroacoustics in the Amundsen Basin near the North Pole (**Figure 46**). On the other hand, the high abundance of amphipods in deep water would not have been revealed without this light attraction as crustaceans were not especially attracted at shallower depths where there probably were relatively less amphipods than other macrozooplankton groups. When only few amphipods were attracted by the light in the lower DSL the availability of amphipods as food for fish was probably very low while it was higher when many amphipods were attracted by the light.

From this study it can be concluded that maximum abundance of macrozooplankton amphipods attracted by red light occurred at 500-700 m of depth, i.e., at the lower end of the DSL, throughout the whole area investigated during the SAS-Oden expedition (between 82.5 °N and 90 °N). It could be hypothesised that for the amphipods their deep occurrence is a strategy to avoid predation by fish in the DSL.

---

[83] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]
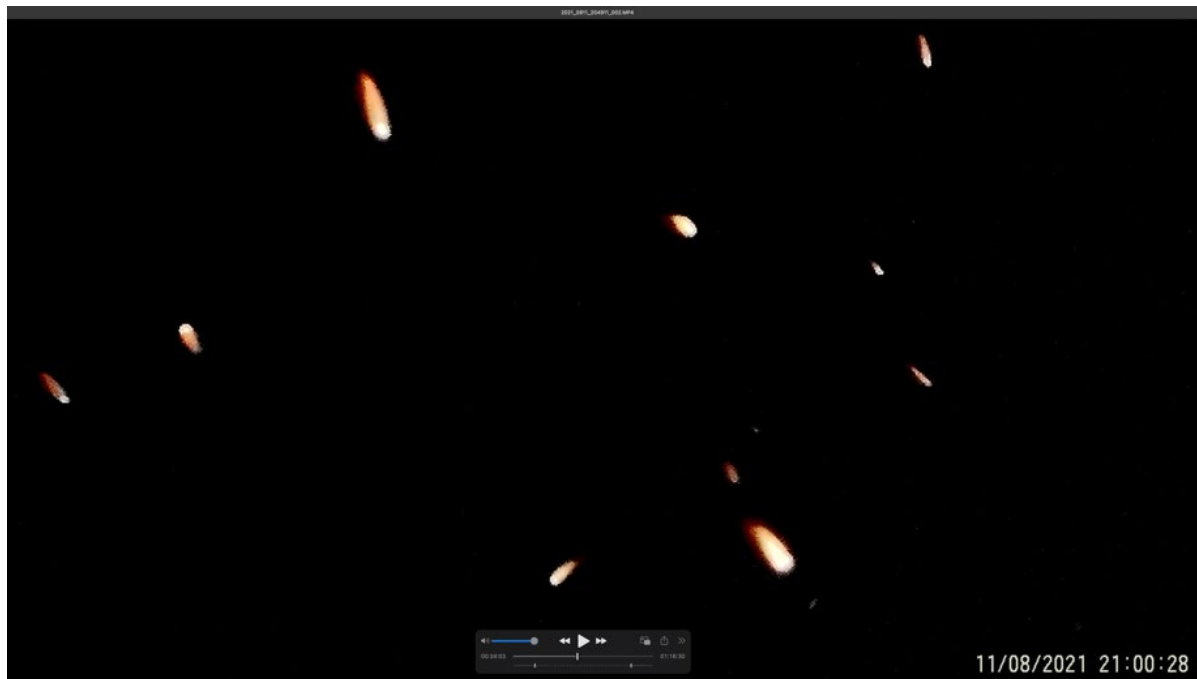
Figure 43: Screen capture taken from the video recording at SAS-Oden Station 14 at 600 m of depth, showing amphipods that are attracted by the red light of the FishCam.

Table 27: Vertical distribution of macrozooplankton individuals (mainly amphipods) analysed from the SAS-Oden FishCam video recordings sorted by latitude (°N). This FishCam was mounted on the CTD and profiles were obtained from 13 stations during the SAS-Oden 2021 expedition. The numbers represent the observed number of macrozooplankton individuals per minute. Red boxes indicate a conspicuous deep-water amphipod maximum, if present. Stn = Station number.

Stn nr	Area	°N	Station Depth (m)	100 m	200 m	300 m	400 m	500 m	600 m	700 m	800 m	Average
58	Yermak Plateau	82.5	-1274	81.3	19.8	31.4	41.1	49.1	53.8	83.1	57.2	52.1
56	Gakkel Ridge	83.9	-2618	35.2	23.1	26.2	17.9	9.8	20.6	11.3	11.5	19.5
50	Greenland shelf	84.2	-889	54.0	13.8	5.4	8.3	7.1	6.6	3.0	8.9	13.4
53	Greenland shelf	84.5	-1351	42.0	12.1	8.3	11.6	11.8	11.1	5.7	11.5	14.3
48	Greenland shelf	84.9	-1554	30.0	4.9	3.6	6.3	6.3	49.8	50.9	7.0	19.8
35	Makarov Basin	87.8	-1389	37.7	13.1	4.4	5.8	16.3	5.6	3.4	7.0	11.7
38	Lomonosov Ridge	87.8	-1198	26.9	16.3	6.8	5.2	11.6	6.8	5.1	3.5	10.3
33	Makarov Basin	88.1	-2987	29.3	10.0	4.6	5.8	8.4	35.6	74.5	31.0	24.9
14	Amundsen Basin	88.5	-4366	87.3	55.0	53.1	18.8	18.5	73.5	18.0	14.6	42.4
30	Makarov Basin	88.6	-3944	27.6	14.8	6.3	4.2	7.5	6.5	4.3	4.4	9.5
26	Lomonosov Ridge	89.1	-1333	26.0	15.9	11.0	9.2	9.6	8.7	8.0	6.3	11.8
18	Amundsen Basin	89.2	-4309	58.9	30.9	27.5	11.6	18.2	23.8	0.0	0.0	21.4
22	North Pole	89.9	-4241	15.9	6.7	6.9	8.4	15.2	45.5	72.8	35.7	25.9
Average 13 stations				42.5	19.2	15.4	12.3	15.8	27.2	26.4	15.8	21.8

Zooplankton distributions derived from hydroacoustic measurements

During the MOSAiC expedition, the ship's echosounder was running almost continuously (see WP2), and thus the acoustic data cover a large temporal range. The zooplankton density detected at 200 kHz showed considerable variability along the MOSAiC drift route (**Figure 44**), which was probably driven by both geographical position and season. During the MOSAiC drift, initially the zooplankton density slightly increased (Leg 1, first half of Leg 2), and then decreased during the second half of Leg 2 while drifting northward in the end of February 2020. The densities continued to decrease until the drifting MOSAiC ice floe reached the Gakkel Ridge and the southern Nansen Basin (Leg 3). Then, a sudden increase was observed, coinciding with increasing influence of Atlantic Water. Water temperature seemed to have a large influence on the zooplankton densities, with the lowest abundances in the coldest water and the highest abundances in the warmest water.

The vertical resolution of the acoustic data showed distinct patterns in the zooplankton distribution (**Figure 45**). In winter the density, expressed as the Nautical Area Scattering Coefficient (NASC), was highest in the upper surface layer (0 m to ca. 50 m). In mid-December, a deeper maximum appears (ca. 150-200 m). This could be due to zooplankton organisms migrating upwards from deeper waters such as lipid-rich copepod *Calanus* species which overwinter at depths below 500 m and move up towards the surface in late winter. However, this conclusion should be verified by comparing the acoustic data with multinet samples. Unfortunately, net data from MOSAiC are not yet available due to time-consuming analyses and lack of man-power. In mid-March, the vertical distribution of the zooplankton patch started to change. The high densities close to the surface shifted downwards, indicate that the previously surface-bound zooplankton migrated to deeper waters. The acoustic data even show that the zooplankton exhibited a well-pronounced diel vertical migration (DVM) during the first two weeks of March. In mid-June, again a surface maximum was observed while densities were low in deeper water. By the end of July throughout August, high zooplankton densities were also found at ca. 150-200 m depth, besides high densities at the surface. In the last weeks of the MOSAiC drift, the surface signal decreased, possibly related to a period of DVM.
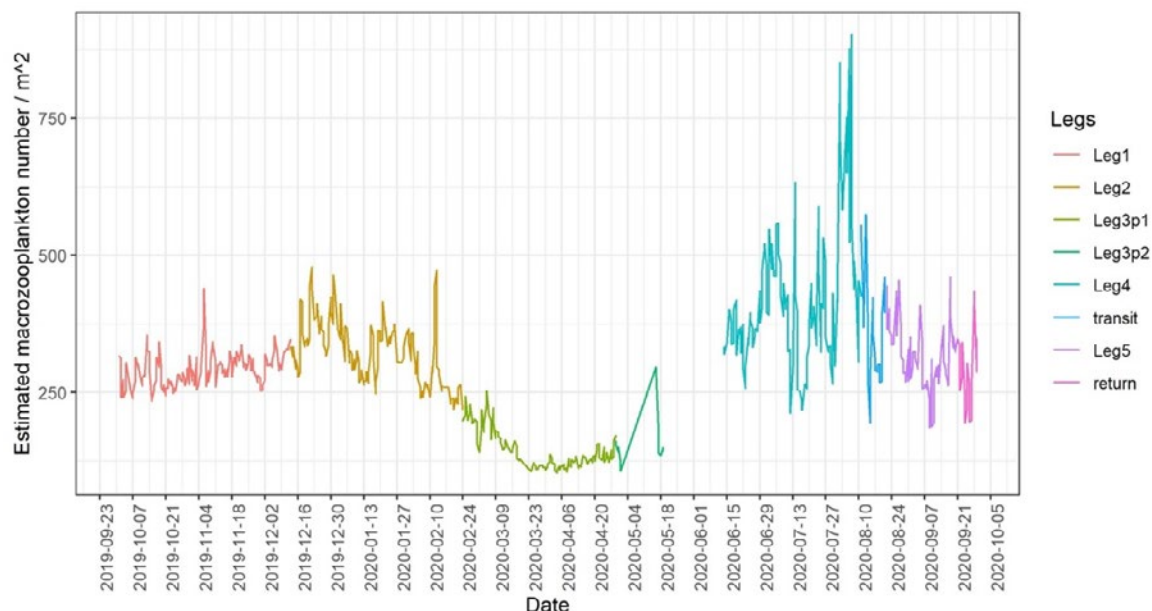
Figure 44: Estimated mean abundance of meso-zooplankton based on 200 kHz acoustic backscatter for the expedition route of MOSAiC, integrated over the upper 150 m of the water column. Averages were calculated for 6-hour intervals.
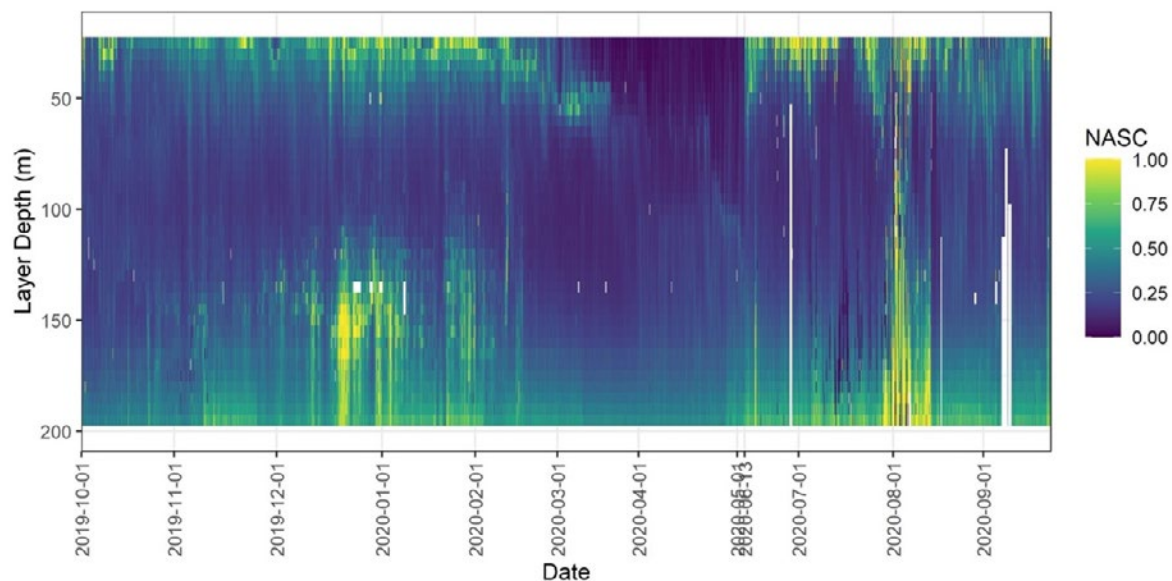
Figure 45: Vertical distribution of mesozooplankton based on 200 kHz acoustic backscatter along the MOSAiC expedition route. Presented are estimated mean abundances expressed as the Nautical Area Scattering Coefficient (NASC, in m^2 nmi^{-2}) of the full water column with 20-m intervals vertically, 1 hour horizontally.

During the SAS-Oden expedition, the WBAT 333 kHz vertical profiles were used to estimate zooplankton biomass distribution in the CAO from the surface to 1000 m depth (**Figure 46**). Data was collected at 19 stations, and abundances in-between were extrapolated. Data were expressed as mean SV which is a relative indicator of the biomass of the zooplankton community.

The zooplankton distribution was characterised by high densities close to the surface followed by a gradual decrease until 200-300 m of depth. Below this, down to 800 m of depth, relatively high zooplankton abundances were detected in the Nansen Basin and the Amundsen Basin. During the north-eastern expedition route, relative zooplankton abundances in deeper water were higher on the Amundsen Basin side but dropped sharply when the Makarov Basin was reached. Thereafter, densities below 400 m in the Makarov Basin, the western Amundsen Basin, the Morris Jesup Rise and the Gakkel Ridge remained very low, suggesting that these regions probably were the most barren regions with respect to zooplankton. At the Yermak Plateau the highest densities in deep water during the entire expedition (NASC>82) were reached.

Based on the detection of fish tracks, it was possible to identify larger organisms from the WBAT 333 kHz vertical profiles collected during the SAS-Oden expedition. It was possible to estimate macrozooplankton abundances, including organisms >2 cm such as amphipods, euphausids, decapods and possibly adults of large copepods such as *Calanus hyperboreus*. Mesozooplankton, however, was too small to being detected by the tracking algorithm.

Comparing macrozooplankton abundances among stations (**Figure 47**) revealed that this group were most frequently detected in the surface layer. In the Nansen and Amundsen Basins, relatively high abundances occurred down to ca. 250-400 m of depth while mesozooplankton was rare below the surface layer of the Makarov Basin. At the Yermak Plateau, besides in the surface layer, relatively high abundances were found also in deeper waters. These data well reflect the overall distribution of zooplankton biomass (**Figure 48**), and it can be concluded that macrozooplankton were of major importance for the biomass distribution pattern, which is not surprising due to their large size compared to mesozooplankton organisms.

The target strength (TS) serves as an indicator for the average size of the macrozooplankton organisms can indicate that different macrozooplankton taxa dominate a zooplankton community. For example, the range between -83 dB and -80 dB is typical for communities dominated by amphipods while communities with an average TS below -85 dB are usually dominated by large copepods such as *Calanus hyperboreus*.

During the SAS-Oden expedition, the mean TS changed considerably with depth. The pattern was similar in the Nansen Basin and the Amundsen Basin (Stations 7, 13, 18, 22) with larger body sizes at the surface and sharply decreasing body sizes below. However, from ca. 50-100 m down to ca. 750 m of depth, body sizes increased to maxima around 750 m of depth in the Nansen Basin) and around 450-500 m of depth in the Amundsen Basin. Below, 750 m body size decreased again. Station 26 on the Lomonosov Ridge showed a different pattern with small-sized macrozooplankton at the surface and two peaks of larger-sized macrozooplankton at ca. 300 and ca. 550 m.

The overall body size of the macrozooplankton was smaller in the Makarov Basin, but the TS depth distribution in principle followed that found in the Amundsen Basin, but with less pronounced second mesopelagic maxima, barely reaching values higher than the -83 dB threshold between *Calanus* and amphipod-dominated communities.

Patterns were less homogenous on the Morris Jesup Rise. At Stations 48 and 50, no clear surface abundance maximum was recorded, and at Station 50 even minimum abundance was detected at the surface, while the body sizes reached maximum values in the surface. At all three stations on the Morys Jesup Rise a second peak was found at ca. 250 m of depth. This peak was distinct at Station 50 and Station 53, but small at Station 48.

At the Gakkel Ridge, the TS was low and comparably stable throughout the entire water column, rarely exceeding -83 dB whereas at the Yermak Plateau the TS was low at the surface and down to ca. 500 m, but it increased to -80 dB at greater depth.

Applying the concept of TS mirroring the dominance of different taxa, i.e., amphipods and *Calanus hyperboreus*, it can be suggested that the macrozooplankton community was often dominated by amphipods in the surface layer while *Calanus hyperboreus* prevailed in most depths below. However, amphipods also dominated the mesopelagic zone, inhabiting depth layers that varied among the stations.
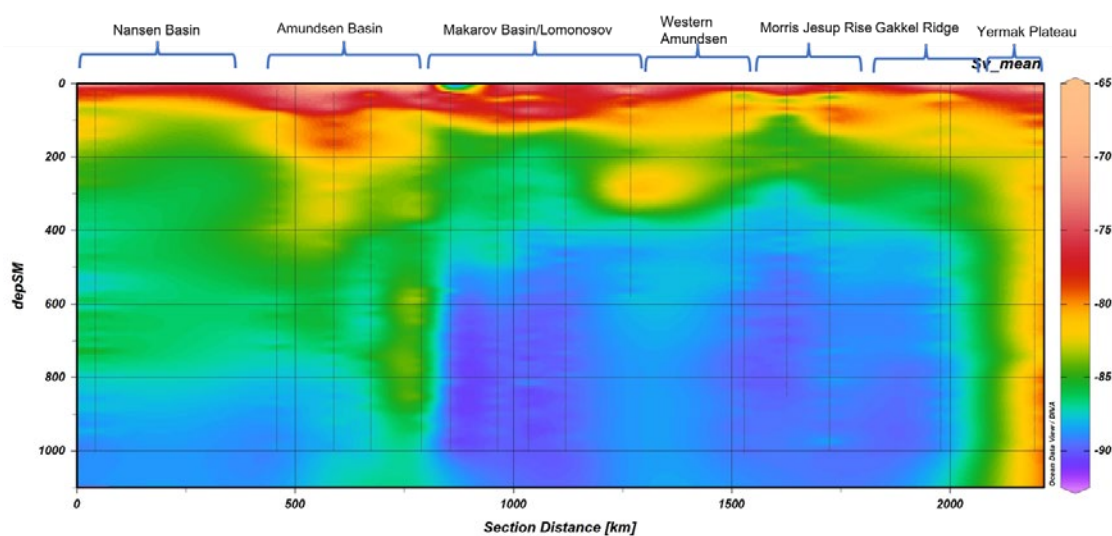
Figure 46: Zooplankton distribution along the SAS-Oden expedition route expressed as mean SV as an index for biomass calculated from hydroacoustic measurements at 19 stations. Vertical lines represent the locations of the stations and sampling depths.
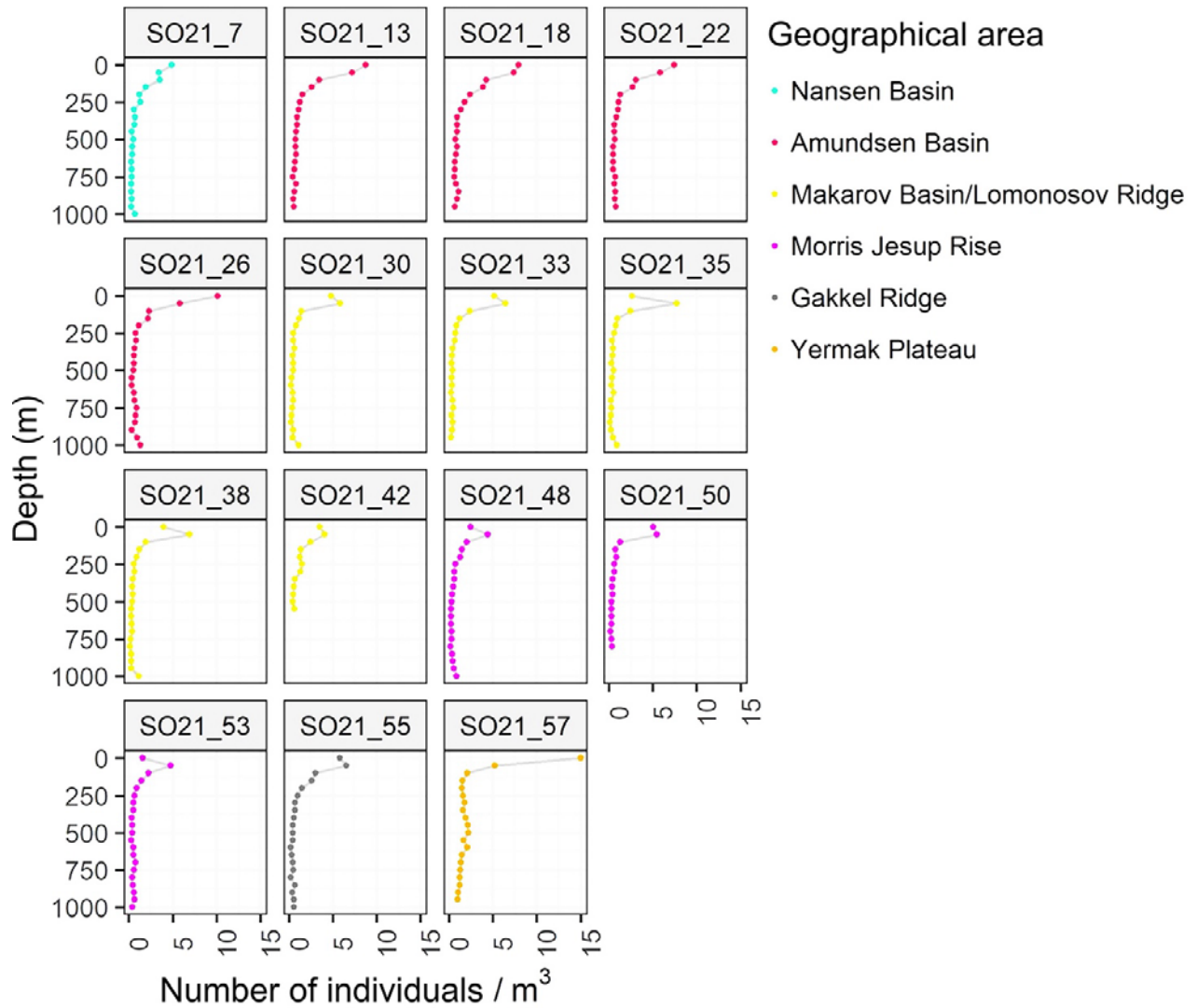
Figure 47: Vertical distribution of macrozooplankton (>2 cm) at 15 stations of the SAS-Oden expedition, estimated from hydroacoustic measurements. The data represent estimated abundances (number of individuals m^{-3}) binned in 50-m intervals of larger zooplankton taxa such as amphipods, euphausiids decapods, and possibly adult Calanus hyperboreus (copepod).

Comparison of zooplankton distributions using different methods

While the methods used for analysing zooplankton during the MOSAiC expedition and the SAS-Oden expedition were the same: multinet mesozooplankton samples (see WP3), LOKI, UVP5 during MOSAiC and UVP6 during SAS-Oden, the status of the sample analyses were different levels at the time this report was submitted. The SAS-Oden samples and data have been completely analysed, but the MOSAiC data (analysed within other research projects) are not yet available due to time-consuming analyses and lack of man-power. Furthermore, for zooplankton acoustics the MOSAiC expedition used a 200 kHz EK80 with continuous measurements and the SAS-Oden expedition used WBAT 333 kHz vertical profiles from CTD casts. However, the combined data allow a comparison among the results from acoustic, optic and net-based approaches, albeit with some limitations.
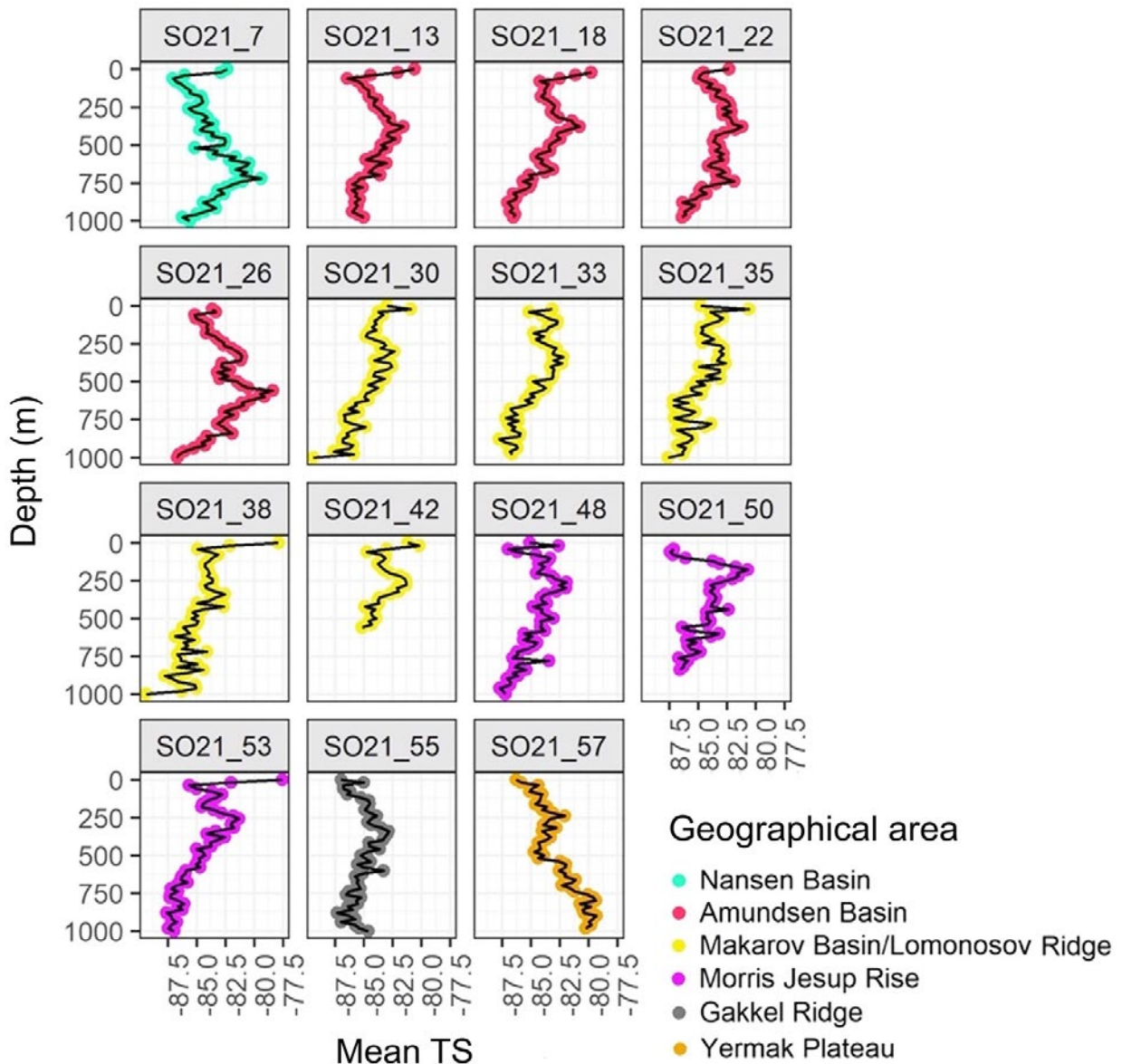
Figure 48: Vertical distribution of mean zooplankton target strength (TS) as indicator of the mean organism size. A range between -83 dB and -80 dB is typical for communities dominated by amphipods. Communities with average TS below -85 dB are usually dominated by large copepods such as *Calanus hyperboreus*.

When comparing mesozooplankton abundances from the SAS-Oden expedition as determined from net samples and optical methods, the most striking, but not surprising, result is that the analysis of multinet samples yielded much higher abundances (**Figure 49 A**) than both the LOKI (**Figure 49 B**) and the UVP6 (**Figure 49 C**). The multinet was on both expeditions equipped with 150 μm mesh size nets, and thus targeting small-sized zooplankton such as the copepod genera *Oithona* and *Microcalanus* well. These genera are abundant in the CAO[84], and *Oithona* usually dominates the community in the upper 50 m while *Microcalanus* resides in high abundances deeper in the epi- and mesopelagic zones (see WP4). The UVP6 only stores images of objects larger than 800 μm and thus severely underestimates the contribution of the smaller mesozooplankton fraction. The images obtained during SAS-Oden suggest that the UVP6 allows for a better

---

[84] Bluhm BA, et al. (2015) A tale of two basins: An integrated physical and biological perspective of the deep Arctic Ocean. Progress in Oceanography 139:89-121 [https://doi.org/10.1016/j.pocean.2015.07.011]

identification of organisms than the UVP5, deployed during MOSAiC. UVP6 images are to date, however, not sufficiently available to draw any final conclusions. The LOKI, on the other hand, is equipped with a 150 μm net as the multinet but, in order to prevent larger organisms from clogging the narrow flow-through chamber, and a 1-mm mesh covers the net opening which leads to escape reactions by fast-swimming organisms.

Also, in terms of general distribution patterns, the methods differed. While multinet and LOKI, both using nets to concentrate the zooplankton organisms, showed higher abundances on the Lomonosov Ridge and low abundances in the Amundsen Basin, relatively more organisms were found in the Amundsen Basin with the UVP. While the differences between UVP6 and LOKI data could be due to the low number of stations sampled in the Amundsen Basin with the LOKI during the SAS-Oden expedition (only one station), the pattern of higher abundances in the Amundsen Basin (7 stations) than in the Makarov Basin (6 stations) determined with the UVP is consistent. It is difficult to compare between abundances that cover such different ranges (multinet on average >100, UVP 8-22 individuals m^{-3} in the upper 1000 m of the water column). However, one explanation could be that small copepods, especially *Microcalanus*, dominated in terms of abundance in the Makarov Basin. Owing to their small size, they are not well, or even not at all, captured with the UVP.

When comparing the results from the SAS-Oden hydroacoustic measurements with the other methods used during this expedition some interesting observations emerge: (1) The zooplankton biomass distribution as recorded with the 333 kHz WBAT vertical profiles (**Figure 45**) and the abundance data derived from the UVP6 casts (**Figure 39**) showed a similar pattern with higher biomasses/abundances in the Amundsen Basin and lower biomasses/abundances in the Makarov Basin. This is remarkable since the WBAT estimated the total zooplankton biomass (both macro- and mesozooplankton) while the UVP6 did not capture macrozooplankton very well and yielded only abundance and not biomass. (2) The ZooScan biomass data based on organism size from the multinet samples (SAS-Oden expedition, unpublished preliminary data) were strongly linearly correlated (r^2=0.817) with the biomass data from the WBAT profiles (calculated separately for the depth intervals sampled by the multinet). Here, again, (mostly) mesozooplankton to (mostly) macrozooplankton biomass were related, and these observations likely reflect a close correlation between meso- and macrozooplankton abundances in the CAO.
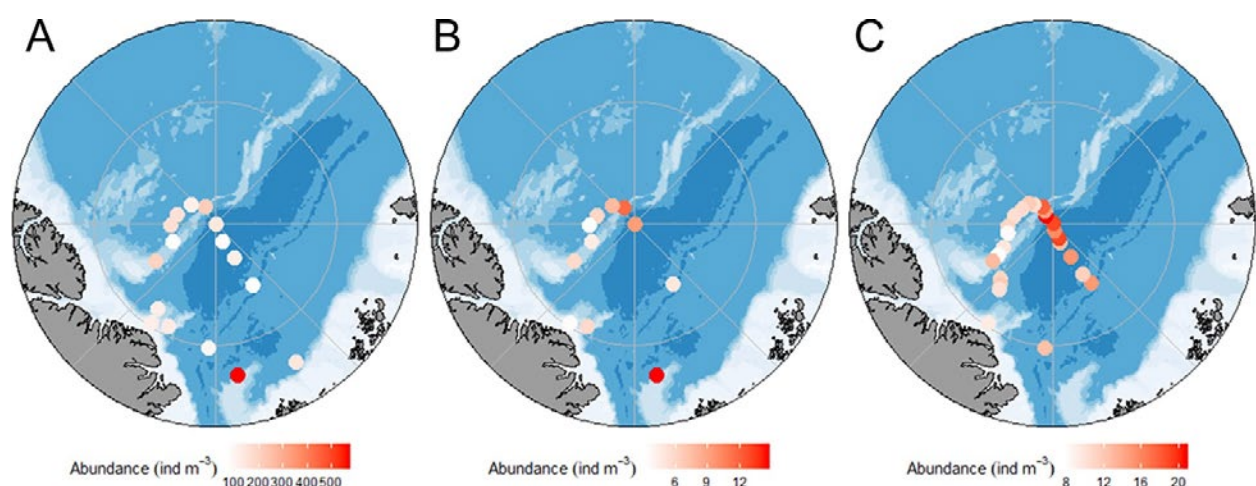
Figure 49: Geographical maps comparing the mesozooplankton distribution along the SAS-Oden route obtained with different methods. (A) Multinet zooplankton samples (SAS-Oden expedition, unpublished preliminary data). (B) Optical data from the LOKI. (C) Optical data from the UVP6. The data represent average abundance expressed as number of individuals per m^{-3} from the surface down to 1000 m of depth. Note the different magnitudes of the scale bars.

Comparison of fish and zooplankton abundance and biomass

Comparing fish and zooplankton biomass on a geographical scale suggests that fish distribution generally mirrors the distribution of their zooplankton prey in the CAO. Zooplankton biomass during the SAS-Oden expedition was higher in the Nansen and the Amundsen Basin, on the Yermak Plateau and on the Lomonosov Ridge than in the Makarov Basin and on the Morris Jesup Rise (**Figure 50**). This general pattern was also reflected in the distribution of fish hydroacoustic backscatter (NASC) and derived fish biomass, except for the Yermak Plateau (WP2). The area near the Lomonosov Ridge where estimated fish biomass was an order of magnitude higher than during the rest of the expedition was also characterised by high mesozooplankton biomass based on the multinet samples (WP4). Furthermore, the hydroacoustic backscatter of macrozooplankton (**Figure 45**) showed comparably high biomass in the deep water layers, even though this is not reflected in the overall zooplankton biomass at some stations.
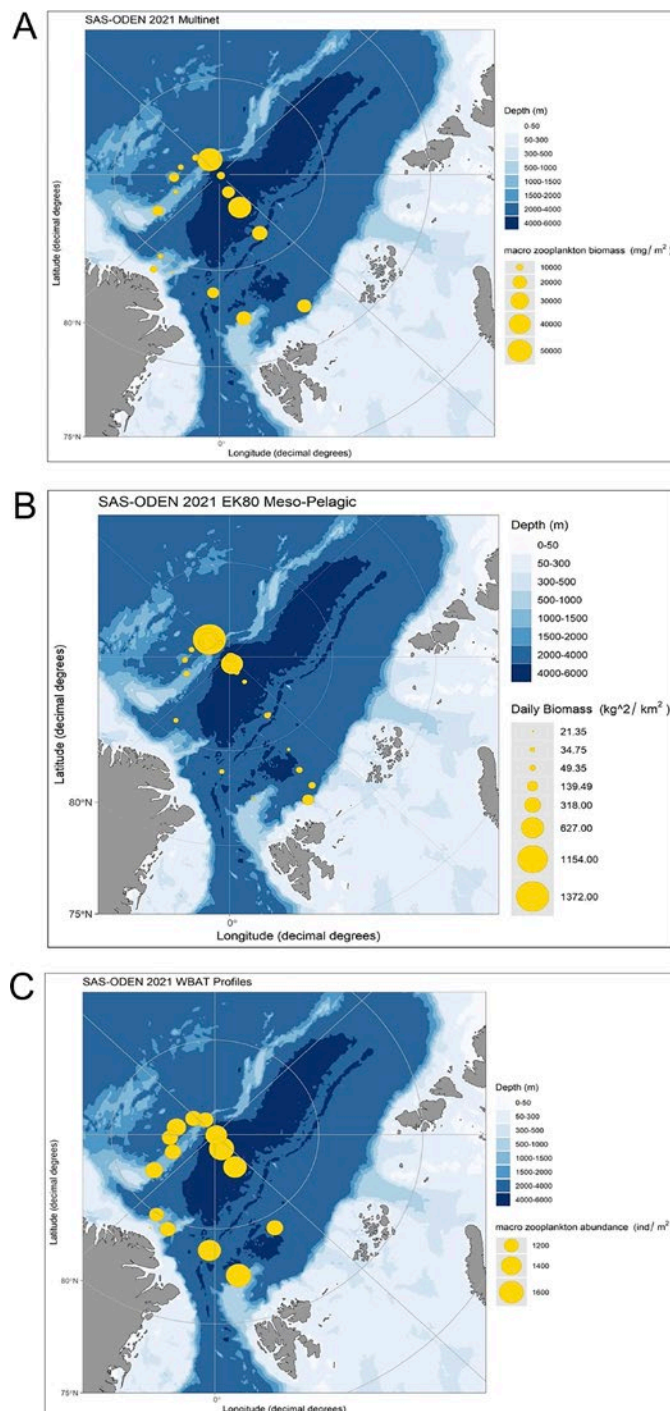
Figure 50: Distribution of zooplankon and fish biomass during the SAS-Oden expedition 2021. (A) mesozooplankton biomass calculated from multinet data (SAS-Oden expedition, unpublished preliminary data). (B) zooplankton biomass calculated from WBAT data (333 KHz). (C) fish biomass calculated from the SAS-Oden EK80 data (18 KHz).

3.6. Answers to the WP3 research questions

(1) When and where did nekton (fish, squid) occur along the MOSAiC and SAS-Oden expedition routes?

Fish were too rare on the videos from the MOSAiC expedition to clearly identify a general occurrence pattern. The armhook squid (*Gonatus fabricii*) was frequently observed from the end of October to mid-December 2019, in the eastern part of the Amundsen Basin. With drifting westward over the Amundsen Basin, the number of observations decreased and the last sighting was on 21 February.

(2) How is the water column structured in terms of particle distribution as a measure of ecosystem productivity?

Particle data from ecosystems north of 80 °N are scarce, and not available for the UVP6. Kiko et al.[85] compiled data from UVP5 casts as available so far from the tropics to polar areas, including both coastal and oceanic regions. When the mean micro-particle concentration in the upper 200 m of the water column (MiP = particles between 102 and 512 μm) were calculated following Kiko et al., the MiP ranged between 0.8 and 15 particles L^{-1} which is at the low end of the 0.8 to 53 486 particles L^{-1} that have been reported by Kiko et al. from all over the world. While the MOSAiC image data are still being analysed, the first results confirm also low particle abundance in the CAO with 0.5 to 30 MiP L^{-1} when RV Polarstern had drifted into the Nansen Basin. Such low particle abundances suggest low productivity. Moreover, the particle sizes, measured during the SAS-Oden expedition were mostly small (<< 512 μm), also suggesting low productivity.

(3) How is the water column structured in terms of zooplankton abundance as possible feeding grounds for fish?

Both, UVP and LOKI show that the largest peaks in mesozooplankton abundance at the very surface (0-20 m) during the SAS-Oden expedition in summer 2021, especially of the lipid rich *Calanus* species clearly visible in the LOKI data. These organisms could thus provide a feeding ground for fish that inhabit the upper water column such as polar cod. Also, a slight increase in abundance in deeper waters (ca. 50 to 250 m was found, depending on the station) which also covered part of the Atlantic Water Layer. Particularly dense mesozooplankton horizons, however, were not found. With regard to macrozooplankton that were important fish prey in the CAO (see WP4 and WP5), the manual MOSAiC FishCam analyses revealed that amphipods were frequent in the deep-scattering layer (DSL)[86] and the manual SAS-Oden FishCam analyses revealed high amphipod abundances in the lower part of the DSL. These results suggest that the macrozooplankton could provide a food source for mesopelagic fish. During the SAS-Oden expedition, acoustic measurements also yielded peaks of large-sized organisms, possibly amphipods, both at the surface and at greater depth. These data thus suggest that there are depth horizons where fish can find favourable feeding conditions.

(4) How does zooplankton abundance estimated by optical and hydroacoustic data relate to net data? (collaboration with WP4)

Hydroacoustic methods mainly target the larger zooplankton group. Unfortunately, the deployments of the ring net that quantitatively captures this size class, were very limited, and the data do not allow for a comparison between the ring net and beam net results. For targeting mesozooplankton, the multinet Midi, UVP and LOKI were deployed.

[85] Kiko R, et al. (2022) A global marine particle size distribution dataset obtained with the Underwater Vision Profiler 5. Earth System Science Data 14:4315–4337 [https://doi.org/10.5194/essd-14-4315-2022]

[86] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

The results from the multinet yielded higher abundances than LOKI and UVP, and is considered the best method to estimate the mesozooplankton biomass available as prey for nekton. Surprisingly, mesozooplankton biomass data as calculated from the multinet samples and macrozooplankton biomass as calculated from acoustic data correlated significantly. This suggests that meso- and macrozooplankton abundances were closely correlated during the SAS-Oden cruise. Future studies are, however, needed to test to which extent mesozooplankton biomass can predict macrozooplankton biomass.

(5) How does fish abundance and biomass relate to zooplankton abundance and biomass? (collaboration with WP2)

The zooplankton biomass during the SAS-Oden expedition was generally higher in the Eurasian Basin, on the Yermak Plateau and on the Lomonosov Ridge than in the Makarov Basin and on the Morris Jesup Rise. This general pattern was also reflected in the distribution of hydroacoustic backscatter and fish biomass except for the Yermak Plateau (see WP2). Remarkably, the area near the Lomonosov Ridge where estimated fish biomass was an order of magnitude higher than during the rest of the expedition was also characterised by relatively high zooplankton biomass. In conclusion, the comparison of the spatial distribution of fish biomass and zooplankton biomass suggests that fish distribution generally mirrors the distribution of their zooplankton prey in the CAO. However, statistical relations between fish and zooplankton were not significant.

3.7. Relevance of the WP3 data for fish stock modelling

With increasing technical progress, imaging and hydroacoustic have become more and more important in pelagic ecosystem research, and mooring optical and acoustical systems in difficult to access areas now provides unique opportunities to study spatial and temporal dynamics of the different compartments of marine ecosystems. The data also impressively show that these approaches yield a yet unpreceded resolution of the vertical and horizontal ecosystem structure, providing knowledge that is essential for fish stock modelling in the CAO.

The MOSAiC FishCam yielded a surprisingly consistent overview of the occurrence of the most abundant cephalopod species in the CAO, the armhook squid (*Gonatus fabricii*). The FishCam, which was moored to the MOSAiC ice flow during the winter months, yielded a surprisingly consistent overview of the occurrence of the most abundant cephalopod species in the CAO, the armhook squid (*Gonatus fabricii*). Extracting images presenting fish was less successful, possibly owing to the low fish abundance but maybe also to the algorithm applied. However, evidence of Atlantic cod in the CAO was found in accordance with the fish catches during MOSAiC. The data derived from the MOSAiC FishCam deployments, however, do not present abundances but number of observations, and can thus not be used for modelling or stock assessments of fish or squid.

Zooplankton is an important component in the Arctic food web, either directly being preyed upon by planktivorous fish or as food for prey of piscivorous fish, and thus not only its geographic distribution but also its vertical location in the water column is important for estimating the prey field for fish. Its overall abundance is low as compared to other (sub)Arctic ecosystems such as the FRAM Strait and the Barents Sea. This is in accordance with the fact that the CAO is an ecosystem with comparably low production[87], as also indicated by low particle number and sizes revealed by the UVP6 deployments during the SAS-Oden expedition.

this study shows that both, meso- and macrozooplankton biomasses may change considerably over very small spatial scales, e.g., in relation to bathymetry when the

---

[87] Bluhm BA, et al. (2015) A tale of two basins: An integrated physical and biological perspective of the deep Arctic Ocean. Progress in Oceanography 139:89-121 [https://doi.org/10.1016/j.pocean.2015.07.011]

Lomonosov Ridge was passed, or in relation to depth. For example, the lipid-rich *Calanus* species were mostly found in the surface and would not be accessible for mesopelagic fish such as Atlantic cod. On other hand, for example amphipods aggregating at greater depth as has been shown by the FishCam video analyses[88] and acoustic data, could provide a rich feeding ground for mesopelagic fish. When *Calanus* migrates to the bathypelagic zone for overwintering, however, it passes the habitat of mesopelagic fish, and could be considered a food pulse. Such dynamics could be vital for the success of a fish population, and simply calculating zooplankton biomass m^{-2} will not reflect the actual prey field. Future monitoring programs should thus include optical systems such as LOKI and UVP, and hydroacoustic methods to determine the fine scale zooplankton abundance of both, macrozooplankton and mesozooplankton. Deploying the UVP, in addition, has the advantage that particle number and size are automatically recorded. These can be used as a proxy for the ecosystem productivity which may change in the years to come due to ocean warming and receding ice cover.

3.8. Recommendations from WP3 for the JPSRM of the CAOFA

Recommended acoustic equipment on the CTD. Back-up equipment is mandatory – motherboards are sensitive and can usually not be repaired on board.
   * WBAT 200 or 333 kHz (zooplankton backscatter)
   * Setup of a very precise accelerometer and compass, or a CTD stabiliser preventing tilt and rotation

Develop a standard "JPSRM set-up" for on-board measurements of zooplankton acoustic properties. These measurements needs to be performed when the zooplankton organisms are alive. Accurate knowledge on macrozooplankton abundance is important for predicting the potential preferred fish habitat.

Deploy "JPSRM ice-tethered autonomous acoustic buoys" that can transmit data to land. This is a relatively cost-effective way to get a lot of acoustic data[89].

Deploy a deep-sea camera system in the DSL. The system should be rated to at least 500 m with electricity and fibre-optic cable, preferably with cameras at different depths in the DSL. This is only worthwhile on drift expeditions / ice stations of at least three weeks due to very low abundances of organisms. Baited camera systems could be tested.

Deploy a Surface and Under-Ice Trawl (SUIT). This trawl was designed for under-ice fish (juvenile polar cod) and ice invertebrates[90].

Design a standard "JPSRM macrozooplankton light-trap line". In the SC07 project it was shown that macrozooplankton species, especially amphipods, are major food items for the mesopelagic fish in the CAO, but these larger animals are rarely caught in multinet samples (targeting mesozooplankton). To estimate access to food for mesopelagic fish, it would be highly recommended to develop a standard light-trap line to be deployed at standard depths in different areas of the CAO for a specific period of time. The macrozooplankton is attracted fast by light (scale of minutes) and deployment time could be one to several hours, but should be tested, The line should cover at least 0-800 m of depth and have weights (to keep the line vertical), a standard-type zooplankton light trap (commercial or especially designed) every 100 m, and a depth/temperature sensor at 400 and 800 m (to check that the line is

[88] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]
[89] Flores H, et al. (on-line manuscript) Sea-ice decline makes zooplankton stay deeper for longer, 09 January 2023, PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-2436026/v1]
[90] Van Franeker JA, et al. (2012) The Surface and Under-Ice Trawl (SUIT). Technical Report https://www.researchgate.net/publication/297794282_The_Surface_and_Under_Ice_Trawl_SUIT#fullTextFileContent

kept vertical). Amphipods were attracted by red light as shown by the FishCam on the CTD during the SAS-Oden expedition, but other colours of light could be tested before defining the standard line that will be used within the JPSRM all over the CAO. This will also provide animals for C:N ratios, stable isotopes and fatty acids analyses from specific depths with higher precision than a MIK net.

Design a standard "JPSRM macrozooplankton sampling net" This should be a closable net so that it is possible to choose a specific depth. During the SAS-Oden expedition a MIK net with a standard diameter of 2 m was used (not closable). This is was large enough to sample macrozooplankton in a representative manner. The usual ring nets (e.g., Nansen net, WP2) are too small, and bigger ones (like the MIK net) do not have a closing mechanism. It should be possible to use a Nansen release mechanism on a large ring net, such as the WP3. Always use standard depths that cover the whole DSL and a bit below, e.g., 1000-500 m, 500-200 m, 200-0 m (to be decided by the JPSRM).

Use a standard "JPSRM mesozooplankton multinet". For example, the much used "Midi". Decide upon standard depth strata: commonly use in the CAO are the depth intervals 2000-1000-500-200-100-0 m (5 nets). Decide upon standard mesh size – commonly used in the CAO is 150 µm. This net targets mesozooplankton, i.e., the prey of smaller fish such as juvenile polar cod and myctophids.

Use standardised image-based measurements for zooplankton. The SC07 project results highlight the potential of computer-aided taxonomic analysis using "ZooScan" in combination with deep-learning software as a time-efficient method to quantify zooplankton species composition, abundance and biomass in the context of a monitoring programme such as the JPSRM. The by the SC07 project established training set on "EcoTaxa" will speed up future development of image analyses of macro- and mesozooplankton of the CAO.

4. FISH PREY AVAILABILITY AND QUALITY (WP4)

4.1. Research questions addressed by WP4

(1) How much fish biomass can be supported by the pelagic ecosystem of the CAO?
(2) What is the quality of zooplankton species as food for fish?
(3) What are the trophic interactions between zooplankton prey species and fish? (collaboration with WP5)
(4) How does zooplankton abundance estimated by net data relate to optical and hydroacoustic data? (collaboration with WP3)

4.2. Data produced by WP4

EFICA_DATA_SC07-WP4.xlsx

For details of the Device Operations (date, time, geographical position, station depth), see files "MOSAiC_Device_operations" and "SAS-Oden_2021_Logbook"[91].

The mesozooplankton multinet data from the MOSAiC expedition are included in a collaboration between the EFICA Consortium and several MOSAiC zooplankton projects coordinated by the AWI and will become publicly available on PANGAEA[92] when analyses are ready.

The mesozooplankton multinet data from the SAS-Oden expedition are included in a collaboration between the EFICA Consortium and one SAS-Oden 2021 project on zooplankton (PI Samuel Hylander, Linnaeus University Kalmar, Sweden)[93], and will become available at the Bolin Centre Database[94] when analyses are ready.

4.3. Human resources of WP4 and main responsibilities

Nicole Hildebrandt (AWI) coordination of the zooplankton analyses, multinet (MN), MIK net, beam net analyses, biomass calculations; Barbara Niehoff (AWI) zooplankton expert advice; Hauke Flores (AWI) stable isotopes (δ^{13}C, δ^{15}N) and C:N ratio analyses and calculations; Pauline Snoeijs Leijonmalm (SU) fatty acids analyses and calculations

4.4. Methods used by WP4

Zooplankton abundance and biomass

Mesozooplankton abundance, community composition and biomass were analysed in 65 samples from vertical MN casts (Hydrobios, Kiel, Germany, net opening 0.25 m^2, a mesh size 150 µm) from 16 stations during the SAS-Oden expedition, covering major realms of the Eurasian part of the CAO (**Figure 51**, **Table 28**). To analyse macrozooplankton abundance and biomass, five samples taken from 800 m depth to the surface with a MIK net (opening 3.14 m^2, mesh size 1.6 mm) were analysed for two stations. Macrozooplankton organisms from these casts were partly sorted on board, identified, counted, photographed and deep-frozen for later fatty acid and δ^{13}C/δ^{15}N analyses while the remaining organisms were preserved in 4% formaldehyde solution. In addition, macrozooplankton were sampled with 42 vertical Beam net casts (ca. 800 m to the surface) at eight stations. All macrozooplankton organisms from these catches were

[91] Bolin Centre Database, Stockholm University [https://bolin.su.se/data/]
[92] PANGAEA Data Publisher for Earth & Environmental Science [https://www.pangaea.de]
[93] Snoeijs-Leijonmalm, P. and the SAS-Oden 2021 Scientific Party (2022). Expedition Report SWEDARCTIC Synoptic Arctic Survey 2021 with icebreaker Oden. Swedish Polar Research Secretariat. 300 pp. [Link]
[94] Bolin Centre Database [https://bolin.su.se/data/]

identified and counted on board for calculating abundance and then frozen at -80 °C for later biochemical analyses. In four casts the samples could not be sorted on board and these samples were preserved in formaldehyde for later analyses.

The formaldehyde-preserved zooplankton samples from MN, MIK and Beam nets were analysed using the digital system ZooSCAN (Biotom, Hydroptic, France)[95,96]. All samples were first size-fractionated (mesozooplankton: 1000/500/70 µm; macrozooplankton: 2000/1000 µm) to reduce the risk of organisms overlapping on the scanner. When zooplankton abundances were high, these subsamples were split into aliquots of up to 1/32 (MN), 1/64 (MIK net) or 1/16 (Beam net) using a Motoda plankton splitter[97]. Then, they were scanned with the ZooSCAN following a standard procedure[98]. Briefly, the scanning area was carefully filled with fresh water, a transparent frame was inserted, air bubbles were removed using forceps, and a background scan was taken using the programme "VueScan" (version 8.3.23). Then, a subsample was transferred into the scanning frame and overlapping organisms were manually separated. The subsample was scanned to obtain a full image with all organisms and particles in the respective sample.

The scans of all aliquots were processed with "ZooProcess" (version 7.19), a macro in ImageJ, to subtract the background, extract images with individual objects, measure each object and link the object with the associated metadata. If there were still several objects on one image, these objects were separated manually in ImageJ. In total, 313 351 images were obtained from the MN samples, 21 595 from the MIK net samples and 3 975 from the beam net samples. All images were uploaded to the "Ecotaxa" web application[99] and automatically assigned to a category using the EcoTaxa classification algorithms based on previously built learning sets (i.e., manually annotated images of Arctic zooplankton) and the Random Forest algorithm provided by Ecotaxa. These predictions were then either validated or manually corrected. When all objects had been annotated, data tables were extracted, including the classification for each object, their size measurements as well as related metadata (station, date/time, location, depth interval, filtrated volume, split factor).

To calculate mesozooplankton abundances (individuals m^{-3}) for the different depth intervals, the counts (n) per subsample and the filtrated volume (V; m^3) as measured by a flowmeter attached to the MN were used:

Abundance = n * split factor / V   (Equation 1)

Biomass (dry weight (DW); mg m^{-3}) was calculated using the maximum (major) and minimum (minor) axes of each organism which were provided by the ZooProcess software. With these two values, the equivalent spherical diameter (ESD; mm) of each individual as an equivalent for body size and the biovolume (BV; mm^3 m^{-3}) were calculated:

ESD = 2 * (major/2 * minor/2)   (Equation 2)

V = 4/3 * п * (major/2*(minor/2)^2) * split factor / V   (Equation 3)

[95] Grosjean P, et al. (2004) Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. ICES Journal of Marine Science 61:518-525 [https://doi.org/10.1016/j.icesjms.2004.03.012]

[96] Gorsky G, et al. (2010) Digital zooplankton image analysis using the ZooScan integrated system. Journal of Plankton Research 32:285-303 [https://doi.org/10.1093/plankt/fbp124]

[97] Motoda S (1959) Devices of simple plankton apparatus. Memoirs of the Faculty of Fisheries, Hokkaido University 7:73-94

[98] Cornils A, et al. (2022) Testing the usefulness of optical data for zooplankton long-term monitoring: Taxonomic composition, abundance, biomass, and size spectra from ZooScan image analysis. Limnology & Oceanography Methods 20:428-450 [https://doi.org/10.1002/lom3.10495]

[99] Picheral M., et al. (2017) EcoTaxa, a tool for the taxonomic classification of images [http://ecotaxa.obs-vlfr.fr]
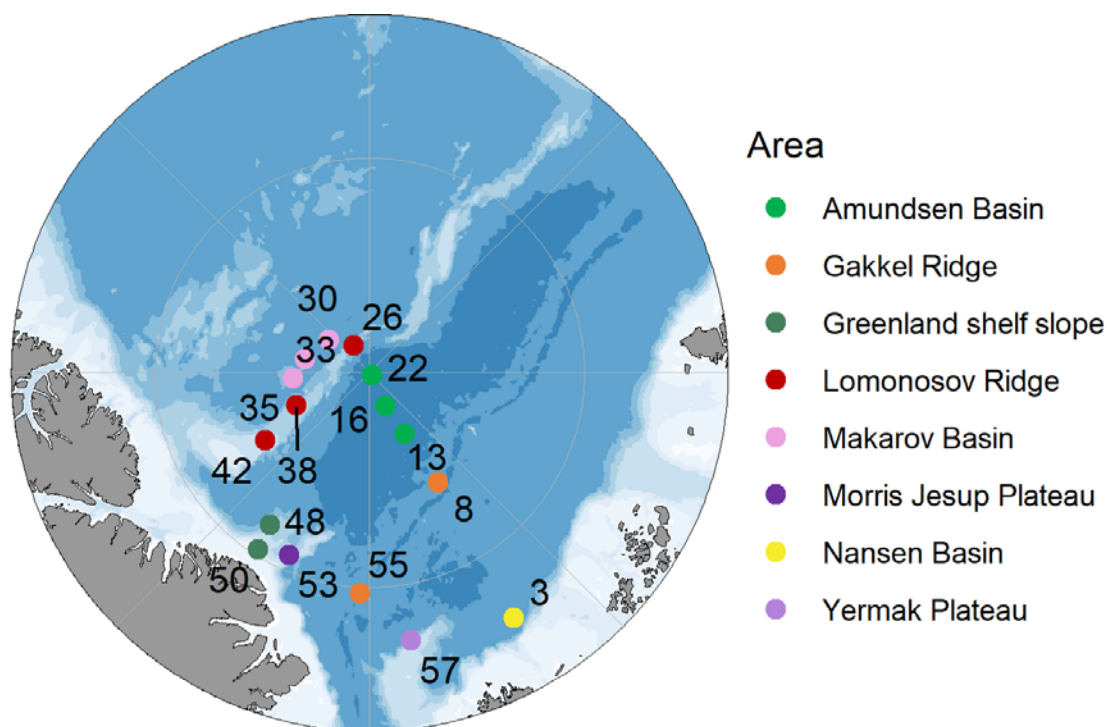
Figure 51: Station map for multinet casts during the SAS-Oden 2021 expedition. Numbers represent station number and colours indicate different geographical areas.

Table 28: List of zooplankton samples taken during the SAS-Oden 2021 expedition. MN = multinet, Beam = beam net, MIK = MIK net.

Station	Date	Area	Latitude (°N)	Longitude (°E)	Bottom depth (m)	Net type (number of samples)
SO21_03	210803	Nansen Basin	82.1	30.5	3224	MN (4)
SO21_08	210808	Gakkel Ridge	86.4	32.1	3037	MN (4)
SO21_13	210810	Amundsen Basin	88.0	29.5	4357	MN (4)
SO21_16	210812	Amundsen Basin	89.0	24.8	4331	MN (2)
SO21_22	210815	Amundsen Basin	89.9	54.3	4136	MN (2), Beam (3)
SO21_26	210819	Lomonosov Ridge	89.1	-149.6	1324	MN (4), Beam (4)
SO21_30	210823	Makarov Basin	88.6	-127.8	3944	MN (4), Beam (5)
SO21_33	210825	Makarov Basin	88.1	-102.0	2960	MN (4)
SO21_35	210826	Lomonosov Ridge	87.8	-86.4	1450	MN (4), Beam (7)
SO21_38	210828	Lomonosov Ridge	87.7	-66.6	1151	MN (4), Beam (7)
SO21_42	210830	Lomonosov Ridge	86.5	-57.3	597	MN (4), Beam (7)
SO21_48	210903	Lomonosov Ridge	84.9	-33.5	1551	MN (4)
SO21_50	210904	Greenland Shelf Slope	84.2	-32.3	901	MN (4), Beam (5)
SO21_53	210905	Morris Jesup Rise	84.4	-23.9	936	MN (4), MIK (2)
SO21_55	210908	Gakkel Ridge	83.8	-2.9	2950	MN (4)
SO21_57	210911	Yermak Plateau	82.4	8.8	1227	MN (4), Beam (4), MIK (3)

The biomass was then calculated by converting the biovolume to wet mass (mg), assuming that the preserved zooplankton organisms are neutrally buoyant with a specific density (ρ) of 1 g cm^{-3}, and then estimating the dry mass (DW) from wet mass by using published conversion factors (CF) for the different taxa[100,101,102,103]:

$$DW = BV * ρ * CF \qquad\qquad \text{(Equation 4)}$$

All calculations were performed in RStudio (version 2022.02.2+485) with the programming language R (version 4.2.0), using scripts that had previously been developed.

To determine macrozooplankton abundances from MIK net and beam net samples, the counts on board were combined with those that were obtained from Ecotaxa annotations. For the MIK net casts, total abundances (individuals m^{-3}) were then calculated according to equation (1). The filtered water volume (m^3) for each cast was in this case calculated from the revolutions (rev) of a mechanical flow meter with back-run stop (Hydrobios, Kiel, Germany) that was attached to the net opening, the pitch of the impeller (0.3 m) and the net opening area (A; m^2):

$$V_{MIK} = rev * 0.3 * A \qquad\qquad \text{(Equation 5)}$$

For the beam net casts only relative abundances could be calculated since the opening area of the net was not fixed and the flow could not be measured.

To calculate size spectra for macrozooplankton organisms, digital images were taken on-board Oden from all sorted animals sampled with either the beam net or the MIK net. To this end, the animals were placed on a plastic plate with a millimetre scale bar. The images were then processed with the software "ImageJ"[104]. First, the tool "Straight line" was used to measure 10 mm on the millimetre scale bar in order to set the scale for the following measurements. Then, the size of the organisms on the image was measured as follows: The length of amphipods, decapods and euphausiaceans was measured from the front of the eye to the end of the telson using a "Segmented line". Chaetognaths were measured from the front of the head to the end of the tail fin, pteropods from the mouth to the tip of the mantle. For cnidarians, the maximum diameter was recorded using the "Straight line". Ctenophores were also measured with a "Straight line", recording the length (i.e., the longest extent) and width. In total, 1 499 individual organisms from beam and MIK net samples were measured. For the remaining individuals, the average lengths of the same species from the same station and sampling device were used, or, if none were available, from all stations and the same device, or, if still no measurements existed, from all measured individuals regardless of the device.

For decapods, euphausiaceans, chaetognaths, the gastropod *Clione limacina* and the amphipod species *Eusirus holmii*, *Lanceola* sp., *Themisto abyssorum*, *Themisto libellula* and *Cyclocaris guilelmi*, biomass was calculated from the length measurements using

[100] Ikeda T & Skjoldal HR (1989) Metabolism and elemental composition of zooplankton from the Barents Sea during early Arctic summer. Marine Biology 100:173-183 [https://link.springer.com/article/10.1007/BF00391956]

[101] Kosobokova K & Hirche HJ (2000) Zooplankton distribution across the Lomonosov Ridge, Arctic Ocean: Species inventory, biomass and vertical structure. Journal of Experimental Marine Biology and Ecology 47:2029-2060 [https://doi.org/10.1016/S0967-0637(00)00015-7]

[102] Postel L, et al. (2000) Biomass and abundance. In: Harris R, et al. (eds.). ICES Zooplankton Methodology Manual. Academic Press: 83-192

[103] Kiørboe T (2013) Zooplankton body composition. Limnology & Oceanography 58:1843-1850 [https://doi.org/10.4319/lo.2013.58.5.1843]

[104] Schneider CA, et al. (2012) NIH image to ImageJ: 25 years of image analysis. Nature Methods 9:671-675 [https://www.nature.com/articles/nmeth.2089]

published length-weight-regressions[105,106,107,108,109]. For *Botrynema* spp. (Cnidaria), a hemispherical shape was assumed to calculate biovolume from the measured diameter (D):

$$BV = (4/3 * \pi * (D/2)^3) / 2 \qquad \text{(Equation 6)}$$

Biomass was then calculated according to Equation 4. For siphonophores and *Aglantha digitale* (Cnidaria), no length measurements could be performed due to their bad visibility on the images. Therefore, the average biovolumes calculated for individuals measured with the Zooscan were applied and then biomass was calculated using conversion factors (Equation 4). The same procedure was applied for the amphipod *Apherusa glacialis*. For the ctenophores, biovolume and biomass were calculated according to Equations 3 and 4.

For the large scyphozoans *Atolla tenella* and *Periphylla periphylla*, no reasonable biomass values could be calculated from measurements of their diameter due to their distinct shape. As they were very rare in the samples (n=24 and n=2 in total, respectively), they were omitted from the abundance and biomass analyses.
Size and species composition data from the beam net and the MIK net were used to calibrate hydroacoustic measurements.

Stable isotopes and C:N ratios

During the SAS-Oden expedition zooplankton animals to analyse C:N ratios and stable isotope analysis of carbon (C) and nitrogen (N) were collected. The animals were first dried for 24 hours in a freeze dryer. After drying, large animals, i.e., fish and macrozooplankton, were ground to powder with piston and mortar, allowing to take a subsample of the tissue. Then, either whole animals (mesozooplankton) or triplicate subsamples of homogenised animal tissue were transferred in a pre-weighed tin cap and weighed to determine the dry mass (DM) of the sample, using a calibrated microscale. The tin caps were then closed and sent to the Institute Littoral Environnement et Sociétés (LIENSS) at La Rochelle, France.

At LIENSS, the samples were analysed with a continuous flow isotope ratio mass spectrometer Delta V Plus with a Conflo IV interface (Thermo Scientific, Bremen, Germany), interfaced with an elemental analyser (EA Isolink, Thermo Scientific, Milan, Italy). The device was calibrated with the following reference material provided by the International Atomic Energy Agency:

$\delta^{13}C$: USGS-24, IAEA-CH6, IAEA-600, USGS-61, USGS-62, USGS-63;
$\delta^{15}N$: IAEA-N2, IAEA-NO-3, IAEA-600, USGS-61, USGS-62, USGS-63.

The analytical precision for these calibration standards was indicated by the manufacturer as 0.10 ‰ for nitrogen and 0.10 ‰ for carbon measurements.

Results were expressed in the δ unit notation as deviations from standards (Vienna Pee Dee Belemnite for $\delta^{13}C$ and N_2 in air for $\delta^{15}N$) following the equation:

[105] Mumm N (1991) Zur sommerlichen Verteilung des Mesozooplanktons im Nansen-Becken, Nordpolarmeer [On the summerly distribution of mesozooplankton in the Nansen Basin, Arctic Ocean]. Berichte zur Polarforschung 92 [https://doi.org/10.2312/BzP_0092_1991]

[106] Böer M, et al. (2005) The Arctic pteropod *Clione limacina*: seasonal lipid dynamics and life-strategy. Marine Biology 147:707-717 [https://doi.org/10.1007/s00227-005-1607-8]

[107] Kreibich T, et al. (2010) Food utilization of two pelagic crustaceans in the Greenland Sea: *Meganyctiphanes norvegica* (Euphausiacea) and *Hymenodora glacialis* (Decapoda, Caridea). Marine Ecology Progress Series 413:105-115 [https://doi.org/10.3354/meps08699]

[108] Kraft A, et al. (2015) Arctic pelagic amphipods: lipid dynamics and life strategy. Journal of Plankton Research 37:790-807 [https://doi.org/10.1093/plankt/fbv052]

[109] Schaafsma FL, et al. (2022) Allometric relationships of ecologically important Antarctic and Arctic zooplankton and fish species. Polar Biology 45:203-224 [https://doi.org/10.1007/s00300-021-02984-4]

$$\delta^{13}\text{C or } \delta^{15}\text{N} = [(R_{sample}/R_{standard}) - 1] \times 10^3 \qquad \text{(Equation 7)}$$

where R is $^{13}\text{C}/^{12}\text{C}$ or $^{15}\text{N}/^{14}\text{N}$, respectively.

For the verification of accuracy and precision, the laboratory standards USGS-61 (caffeine) and USGS-63 (caffeine) were analysed three times per measurement batch of maximum 96 samples. The standards were analysed in duplicates. The analytical standard deviations were 0.04 ‰ for $\delta^{13}\text{C}$ in both standards, and 0.04 and 0.05 for $\delta^{15}\text{N}$ in USGS-61 and USGS-63, respectively.

The analytical results of the mass spectrometer also include the relative mass percentages of C and N in each sample. The C and N content of each sample was calculated according to the equation:

$$C \text{ or } N = \frac{\%C \text{ or } \%N}{100} \cdot M_s \qquad \text{(Equation 8)}$$

where C and N are the mass (in mg) of C and N, respectively, %C and %N are the relative mass percentages of C and N, respectively, and M_s is the dry mass of the sample in mg. The C:N ratio of each sample was then calculated by dividing C by N:

$$C{:}N = \frac{C}{N} \qquad \text{(Equation 9)}$$

Fatty acid analyses

The fatty acid (FA) composition can be used as a trophic biomarker for food-web interactions between zooplankton and fish, since certain essential FAs are incorporated unmodified from the diet of an organism and the relative contribution of FAs changes with the trophic level[110]. This implies that the FA composition of a zooplankton organism reflects that of its diet and that of a fish reflects that of its zooplankton prey (**Table 29**). In the open ocean only phytoplankton, for example, synthesise long-chain polyunsaturated essential FAs (PUFAs), such as 20:5n-3 and 22:6n-3, which are transferred to higher trophic levels, and are thus the two predominant PUFAs in aquatic food webs.

Besides revealing trophic interactions, the FA content is also a measure for food quality. Prey with high FA levels and high proportions of PUFAs are considered favourable food because these FAs are essential components for animal growth. The high energy potential of monounsaturated fatty acids (MUFA) is of particular importance in cold waters, and various fatty acids including saturated fatty acids (SAFA) are involved in organismal responses to environmental stressors.

During the SAS-Oden expedition, individual zooplankton were collected alive from net samples, transferred to precombusted Wheaton vials, and immediately deep-frozen at -80 °C. In total, 50 samples with single zooplankton individuals, but for the two smaller copepod species *Calanus glacialis* and *Metridia longa* 6-10 individuals were analysed per sample. These included four decapods (all *Hymenodora glacialis*), 20 amphipod (six taxa), 14 copepod (four taxa), three chaetognath (one taxon) and nine jellyfish (three taxa) samples. In the home laboratory, the samples were removed from the freezer, and each individual was weighed (g, with five decimals), placed into a separate Eppendorf tube. The whole animal bodies were freeze-dried and homogenised before lipid extraction.

The fatty acid contents and compositions were analysed by gas chromatography – mass spectrometry (GC-MS) at the Swedish Metabolomics Centre (SMC) in Umeå. For the

---

[110] Dalsgaard J, et al. (2003) Fatty acid trophic markers in the pelagic marine environment. Advances in Marine Biology 46:225-340 [https://doi.org/10.1016/s0065-2881(03)46005-7]

quantification of FAs bound to other molecules, mainly in cell membranes, a sub-sample was transmethylated. For the quantification of free FAs, another sub-sample was methylated. One µl of each subsample was injected in split mode 1/10 into an Agilent 7890A GC by an Agilent 7693 autosampler. The GC was equipped with a multimode inlet (MMI) and with a Zebron ZB-FAME 20 m × 0.18 mm in diameter fused silica capillary column with a chemically bonded 0.15 µm stationary phase (Phenomenex, Torrance, USA). The GC and MS settings and the column ramping schedule were according to the SMC standard procedure. The column effluent was introduced into the electron impact (EI) ion source of an Agilent 7000C QQQ mass spectrometer where the ions were generated by a 70 eV electron beam at an emission current of 35 µA and analysed in dMRM-mode. The solvent delay was set to 2 minutes. This method was able to separate 18:1n-9, trans and 18:1n-9, cis, but not 18:1n-7 and 18:1n-7. The latter two compounds therefore can be mixed with either 18:1n-9, trans or 18:1n-9 or both.

Table 29: Trophic markers and ratios commonly analysed in fatty acid profiles. Table from Kraft et al.[111]

Fatty acid	Trophic marker
16:0	Carnivory
16:1(n-7)	Spring bloom (diatoms)
16:4(n-1)	Diatoms / ice algae
18:1(n-9)	Carnivory
18:2(n-6)	Chlorophytes or cyanobacteria
20:1(n-9)	*Calanus* spp.
20:5(n-3)	Diatoms
22:1(n-11)	*Calanus* spp.
22:6(n-3)	Flagellates, e.g., the presence of *Phaeocystis pouchetii* in the diet of *Calanus*
Fatty acid ratios	**Trophic marker**
18:1(n-9) / 18:1(n-7)	High values (>3) as indicator for decreasing carnivory in marine zooplankton
20:5(n-3) / 22:6(n-3)	High ratio – diatom-originated diet; low ratio – flagellate-based diet
PUFA / SFA	Increasing value may be used as an indicator for dominance of carnivorous versus herbivorous feeding; however also increases under starvation conditions

4.5. Results and discussion of WP4

Mesozooplankton biomass

During the SAS-Oden expedition, the overall mesozooplankton abundance, averaged over the entire sampling depth from 1000 m (or bottom) depth, was low throughout the study area, ranging from 69 to 217 individuals m^{-3} (**Figure 52 A**). Only at station 57, on the Yermak Plateau, higher abundances of 577 individuals m^{-3} were found. The total zooplankton biomass ranged from 4.1 to 20.9 mg dry weight m^{-3} and the spatial pattern differed slightly from that of zooplankton abundances (**Figure 52 B**). The highest biomass was found on the north-east transect from the Nansen Basin across the Gakkel Ridge and the Amundsen Basin towards the Lomonosov Ridge as well as on the Yermak Plateau. In the Makarov Basin and on the Greenland shelf the zooplankton biomass was relatively low.

The data are consistent with the low zooplankton abundances and biomasses that have been reported from the CAO earlier[112]. Higher biomass has been shown to be associated with the Atlantic water inflow, and in consistence with the hydrographic circulation pattern, the highest biomass during the SAS-Oden expedition were found in the north-

[111] Kraft A, et al. (2015) Arctic pelagic amphipods: lipid dynamics and life strategy. Journal of Plankton Research 37:790-807 [https://doi.org/10.1093/plankt/fbv052]

[112] Bluhm BA, et al. (2015) A tale of two basins: An integrated physical and biological perspective of the deep Arctic Ocean. Progress in Oceanography 139:89-121 [https://doi.org/10.1016/j.pocean.2015.07.011]

eastern part of the study area. Although not all zooplankton samples that were collected during the MOSAiC expedition have yet been analysed, preliminary data show that the highest abundances at stations most heavily influenced by Atlantic water, i.e., in the Nansen Basin and the Fram Strait, and thus support the results from the SAS-Oden expedition.
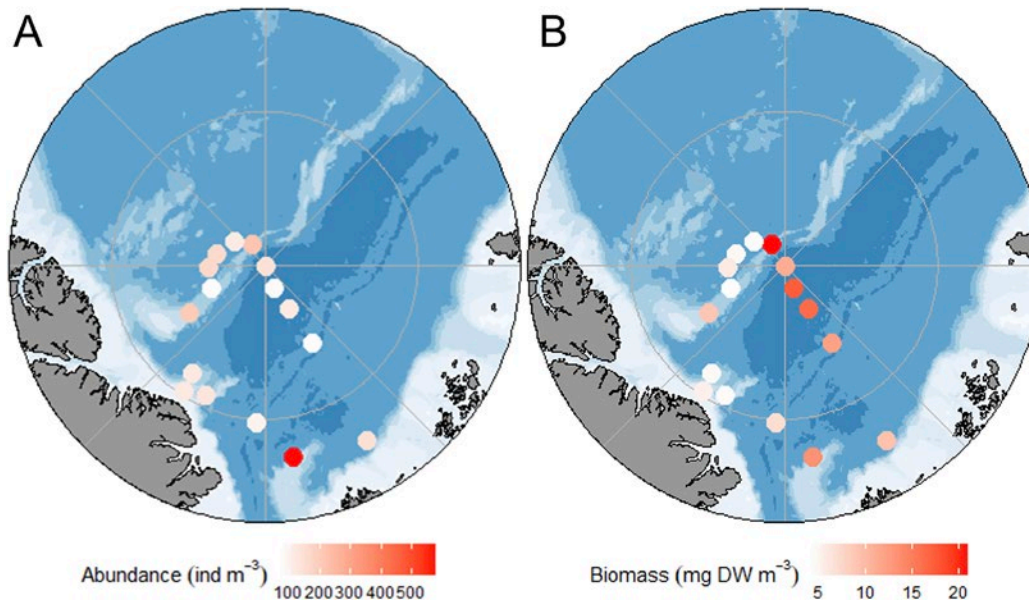
Figure 52: Mesozooplankton abundance and biomass in the upper 1000 m of the water column obtained by image analysis of samples from multinet casts during the SAS-Oden expedition. (A) average abundance. (B) average biomass. DW = dry weight

Mesozooplankton community composition

Mesozooplankton community composition did not vary much along the SAS-Oden expedition route for higher taxonomic levels. Crustaceans dominated the communities in terms of abundance (relative abundance >73%) from the surface down to 1000 m depth at all stations, except for Station 57 (0-50 m) on the Yermak Plateau where appendicularians were the most abundant group in the surface layer. Appendicularians was an abundant group at many stations (5-8%), especially towards the end of the expedition. It is possible that this was due to the seasonal succession rather than to spatial differences as appendicularians form blooms later in the year. All other higher taxonomic groups were much less abundant with some exceptions: gastropods relative abundances were 3-7% in the Makarov Basin and at one station on the Lomonosov Ridge and foraminifers contributed with up to 15% to the community in the subsurface layer (50-200 m) at Stations 3, 8 and 13 (Nansen Basin, Gakkel Ridge and Amundsen Basin, respectively). For all other taxa relative abundances were below 5%.

Macrozooplankton

Macrozooplankton dominated in the fish stomachs analysed (see WP5) and was sampled at two Stations the Morris Jesup Rise and the Yermak Plateau with a MIK net. Comparing these two stations, abundance was low (<0.2 individuals m^{-3}) at Station 53 (Morris Jesup Rise), while they varied between 0.9 and 2.2 individuals m^{-3} among the three casts at Station 57 (Yermak Plateau). In accordance, also the biomass were higher on the Yermak Plateau than on the Morris Jesup Rise (**Figure 53 A**).
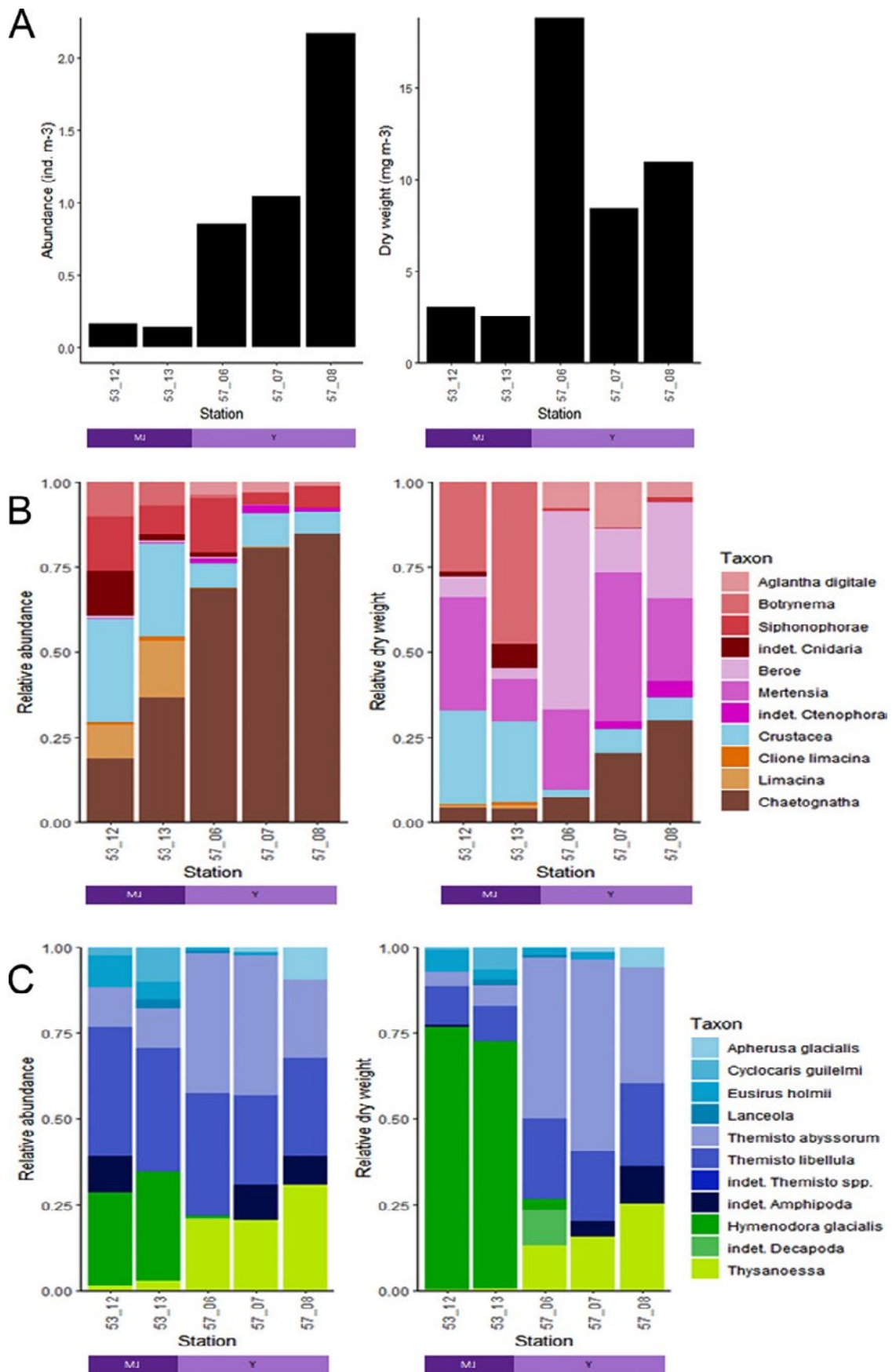
Figure 53: Macrozooplankton abundance, biomass and community composition at two stations in the Central Arctic Ocean (CAO) during the SAS–Oden expedition analysed from five MIK net tows. (A) Total abundance and biomass. (B) Relative abundances in the total macrozooplankton community. (C) Relative anbundances of crustaceans only.

On the Morris Jesup Rise (Station 53), community composition was well-balanced in terms of abundance. Cnidarians (mostly hydrozoans), crustaceans, chaetognaths and molluscs contributed almost equal parts to the community (**Figure 53 B**). Among the cnidarians, *Aglantha digitale*, *Botrynema* spp. and siphonophores were identified. Siphonophora was the most abundant taxon while *Botrynema* spp. were most important in terms of biomass. Ctenophores, including the genera *Beroe* and *Mertensia*, were not abundant at this station, but the few individuals, especially *Mertensia* specimens, substantially contributed to the total community biomass. Molluscs were mostly represented by *Limacina* spp., a shelled pteropod, and its predator, the sea angel *Clione limacina*. Among the crustaceans, the decapod species *Hymenodora glacialis*, contributed most to the biomass whereas several amphipod species were dominant in terms of abundance (ca. 70%), but not in biomass (ca. 25%) (**Figure 53 C**).

On the Yermak Plateau (Station 57), chaetognaths contributed more to the macrozooplankton community in terms of abundance in all three casts than any other taxon, but their biomass contribution was relatively low (<30%; **Figure 53 B**). Cnidarians and crustaceans were the second-most abundant groups, but not in terms of biomass. Molluscs were almost absent at this Station 57. Most of the biomass at this station consisted of ctenophores. The same ctenophore genera as on the Morris Jesup Rise were present (*Beroe* and *Mertensia*), but *Beroe* contributed more to the biomass on the Yermak Plateau than on the Morris Jesup Rise. In contrast to the Morris Jesup Rise, *Thysanoessa* spp. (Euphausiacea) were relatively abundant within the crustaceans, both in terms of abundance and biomass, and while the relative amphipod biomass was rather low the Morris Jesup Rise, it was high (>74%) on the Yermak Plateau (**Figure 53 C**).

The vertically towed beam net was deployed 42 times during the SAS-Oden expedition. This net was designed to in the first place catch fish and therefore it had to be large. Initially, a 10-m diameter ring net was constructed, but when this net was tested in the Baltic Sea it appeared that the construction was not strong enough to withstand the pressure from fishing down to 800 m of depth. Instead, the beam net, a net with a 10-m long steel beam in the middle to achieve maximum opening in the cross-track direction and two aluminium otter boards (trawl doors) of each 1.5 m^2 and kites attached to the headrope to open the net was designed[113]. With this beam net no fish was caught during the SAS-Oden expedition (except for one sympagic polar cod), only macrozooplankton. The net opening of the beam net was not always opened maximally by the kites but the opening varied with the prevailing water currents (as opposed to a ring net with a fixed diameter), and the water volume passing through the net could not be quantified. Therefore, it was impossible to achieve robust quantitative estimations of the macrozooplankton (individuals m$^{-3)}$ and relative abundance estimates were used as a rough measure of community composition, expressed as number of organisms m^{-1} simply calculated from the distance the net was towed through the water, and for the ease of comparison, abundances and biomasses from all casts performed at one station were averaged.

Abundance (individuals m^{-1}) was highest on the Yermak Plateau (Station 57) (**Figure 54 A**). The total macrozooplankton biomass differed between stations, with the lowest value in the Amundsen Basin (Station 22) and highest values along the Lomonosov Ridge (Stations 35, 38, 42). The community composition of the beam net samples was quite similar to that from the MIK net, but with higher relative abundances of crustaceans. Crustaceans dominated the communities in terms of abundance at Stations 22, 26 and 30, while cnidarians dominated at Station 35, 38, 42 and 50 and chaetognaths were most abundant at Station 57 (**Figure 54 B**). Among the non-crustaceans, the hydrozoa *Botrynema* spp. was relatively abundant and also provided considerable biomass at all stations, except for Station 57 on the Yermak Plateau. Also, the ctenophore *Beroe* was important in terms of biomass, but not in terms of abundance. The hydrozoan *Aglantha digitale* only occurred at Station 57.

---

[113] Snoeijs-Leijonmalm P, et al. (2021) Ecosystem mapping in the Central Arctic Ocean (CAO) during the MOSAiC Expedition. Publications Office of the European Union [https://data.europa.eu/doi/10.2926/714618]
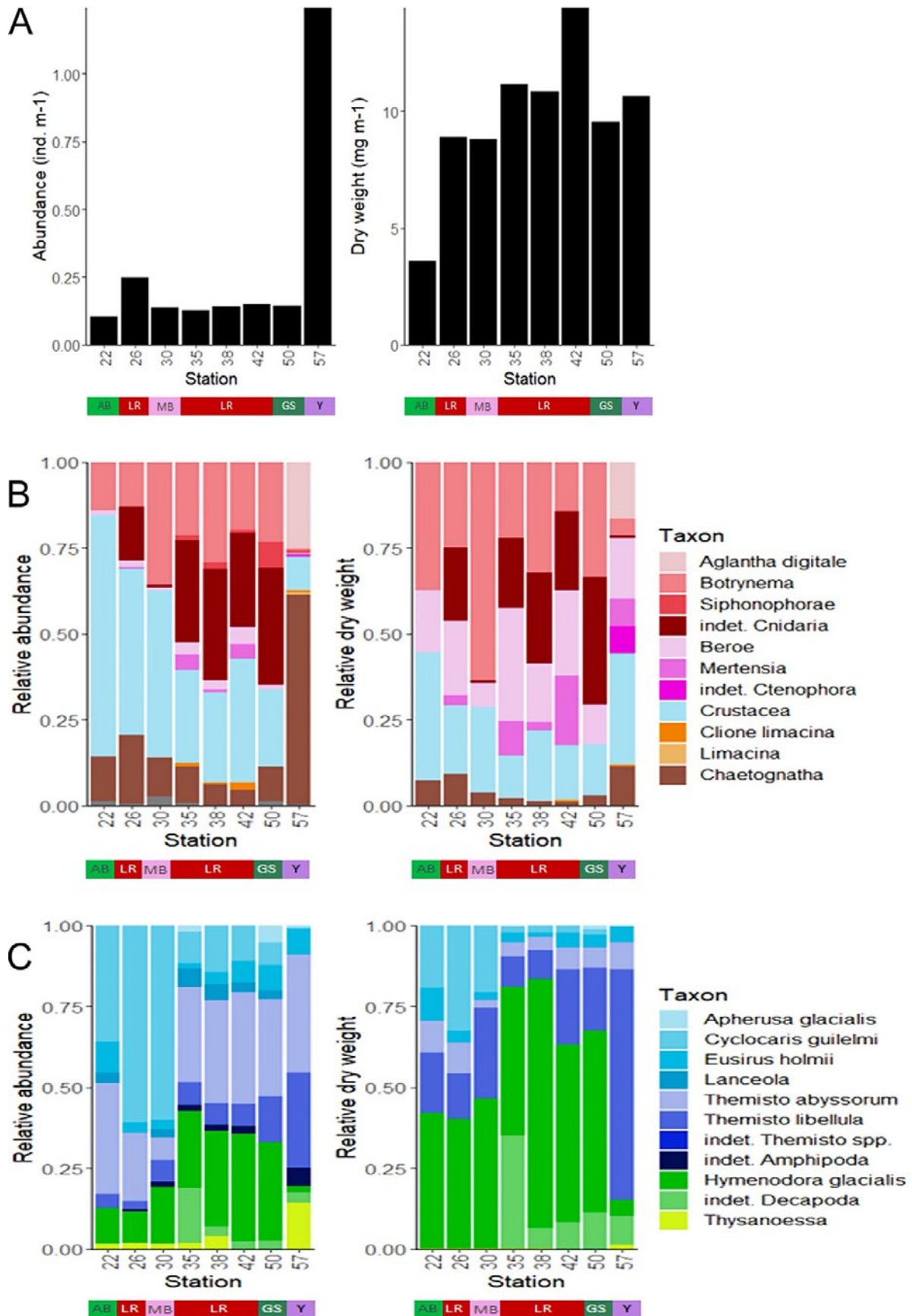
Figure 54: Macrozooplankton community composition at eight stations in the Central Arctic Ocean (CAO) analysed from 42 beam net tows. (A) Total abundance and biomass. (B) Relative abundances in the total macrozooplankton community. (C) Relative anbundances of crustaceans only.

Among the crustaceans (**Figure 54 C**), the amphipod species *Cyclocaris guilelmi* dominated at Stations 22, 26 and 30 (Amundsen Basin, Lomonosov Ridge and Makarov Basin, respectively) in terms of abundance, and this species also considerably contributed to the biomass to the macrozooplankton community. In contrast, *Themisto abyssorum* (Amphipoda) and Decapoda (*Hymenodora glacialis* and unidentified decapods) were more abundant at Stations 35, 38, 42 and 50 (Lomonosov Ridge and Greenland Shelf Slope). On the Yermak Plateau (Station 57), both the relative abundance and the biomass of the amphipod *Themisto libellula* were high. At all other stations, the bulk of the crustacean biomass consisted of the decapod *Hymenodora glacialis*.

Zooplankton food web structure based on δ¹³C and δ¹⁵N and C:N ratio

The stable isotope composition and C:N values of 41 *Cyclocaris guilelmi*, seven *Hymenodora glacialis*, nine *Lanceola clausi*, 48 *Themisto abyssorum*, and 53 *Themisto libellula* sampled during the SAS-Oden expedition were analysed (**Table 30**; **Figure 55**). The ranges of the measured values were well within the ranges of polar cod sampled during both expeditions (see WP5), and was similar to the ranges for the respective zooplankton species from the MOSAiC expedition (K. Schmidt, personal communication).

These data suggest that *Cyclocaris guilelmi* is a predatory amphipod feeding on the partly ice algae-driven food web of the CAO. *Hymenodora glacialis*, *Lanceola clausi*, *Themisto abyssorum* and *Themisto libellula* were in the same range of δ¹³C values, also suggesting some influence of ice-algae in their diet. The δ¹⁵N values in *Themisto libellula* and a group of *Themisto abyssorum* individuals, however, were lower than in *Cyclocaris guilelmi*, *Hymenodora glacialis* and *Lanceola clausi* with δ¹⁵N values between 7 and 13, indicating that they were feeding about one trophic level below (**Figure 55**).

The mean C:N ratios were 6.3 in *Cyclocaris guilelmi,* 7.8 in *Hymenodora glacialis*, 5.8 in *Lanceola clausi*, 5.3 in *Themisto abyssorum* and 6.7 in *Themisto libellula*, implying a relatively high nutritious value of these species (**Table 30**). These crustaceans were important prey items in Atlantic cod caught in the CAO (see WP5).

Table 30: δ¹³C, δ¹⁵N and C:N ratio from the main zooplankton species for fish sampled during the SAS-Oden 2021 expedition. The data are shown as mean ± 1 standard deviation. n = number of animal individuals analysed.

Species	n	δ¹³C (‰)		δ¹⁵N (‰)		C:N	
		Mean	Range (min to max)	Mean	Range (min to max)	Mean	Range (min to max)
Cyclocaris guilelmi	41	-25.3 ± 1.6	-28.8 to -22.8	14.4 ± 1.1	10.0 to 16.0	6.3 ± 2.2	4.0 to 13.4
Hymenodora glacialis	7	-25.7 ± 1.2	-27.2 to -24.1	12.9 ± 1.7	11.1 to 15.0	7.8 ± 2.8	4.0 to 11.1
Lanceola clausi	9	-26.4 ± 0.7	-27-6 to -25.5	13.1 ± 1.7	11.8 to 17-4	5.8 ± 1.2	4.5 to 8.3
Themisto abyssorum	48	-25.7 ± 1.6	-30.0 to -23.3	12.1 ± 2.2	7.1 to 15.1	5.3 ± 1.4	3.9 to 9.3
Themisto libellula	53	-26.8 ± 1.3	-29.2 to -23.5	10.7 ± 1.8	5.8 to 13.2	6.7 ± 1.5	4.1 to 10.2
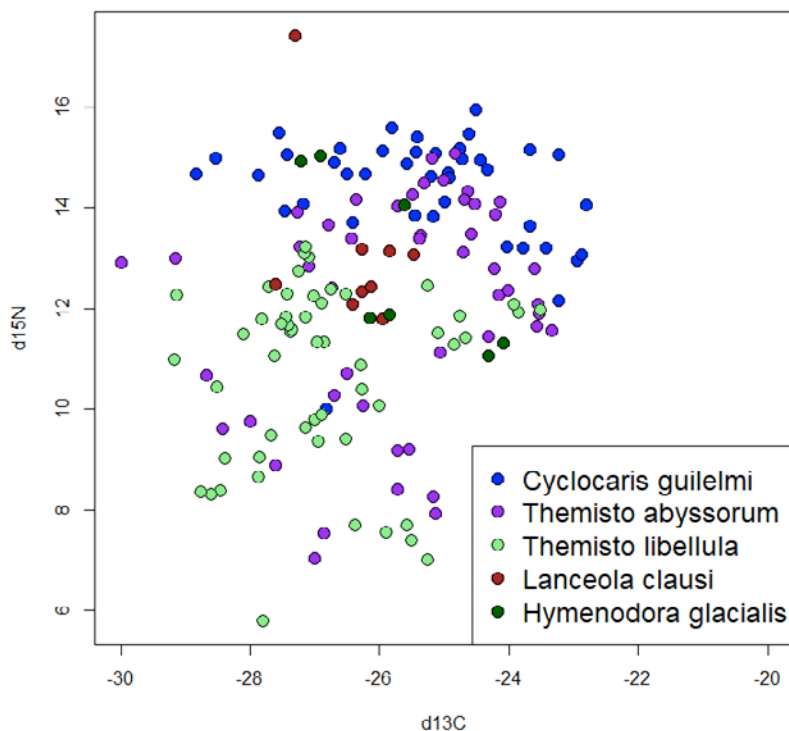
Figure 55: Plot of δ¹⁵N versus δ¹³C in the main zooplankton prey species for fish sampled during the SAS-Oden 2021 expedition.

Zooplankton food-web structure based on fatty acid composition

The total FA concentrations in ng per mg wet weight were highest in the crustaceans and the lowest in the jellyfish (**Figure 56**), indicating that jellyfish have low nutritious value per unit biomass ingested. In all crustaceans the concentrations of bound FAs were 2-8 times higher than those of free FAs, while in jellyfish concentrations of bound FAs were even up to 22 times higher. However, this does not affect their nutritional value for fish.

As expected, the comparably large, lipid-storing copepod *Calanus glacialis* had the highest specific FA content per mg ww^{-1}. In *Calanus hyperboreus*, which is one of the largest copepods worldwide and also stores large amounts of lipids, the specific FA content was, however, surprisingly low, and did not reach the values of the other two copepod taxa *Metridia longa* and *Paraeuchaeta* spp.. This result cannot be fully explained, but it is possible that this is due to the developmental stages that have been sampled. The data on *Calanus glacialis* reflect the lipid content of females that were possibly well-fed and ready to spawn as this species reproduces during the summer in the surface layer[114]. In contrast, *Calanus hyperboreus* spawns in deeper water in winter and spring, fuelled solely by internal lipid reserves, and thus at the end of the reproductive season females are poor in lipids. Copepodite stage IV of *Calanus hyperboreus* did not contain as many lipids as the subadult stage CV or adult females prior to spawning. The overall low fatty acid content may indicate that under the thick ice cover encountered during the SAS-Oden expedition, phytoplankton production was still insufficient to replenish the lipid stores of *Calanus hyperboreus.* In contrast, *Calanus glacialis* are able to exploit ice algae which bloom prior to the phytoplankton, and hence were already accumulating lipids for egg production.

---

[114] Niehoff B (2007) Life history strategies in zooplankton communities: The significance of female gonad morphology and maturation types for the reproductive biology of marine calanoid copepods. Progress in Oceanography 74:1-47 [https://www.sciencedirect.com/science/article/abs/pii/S0079661107000432]
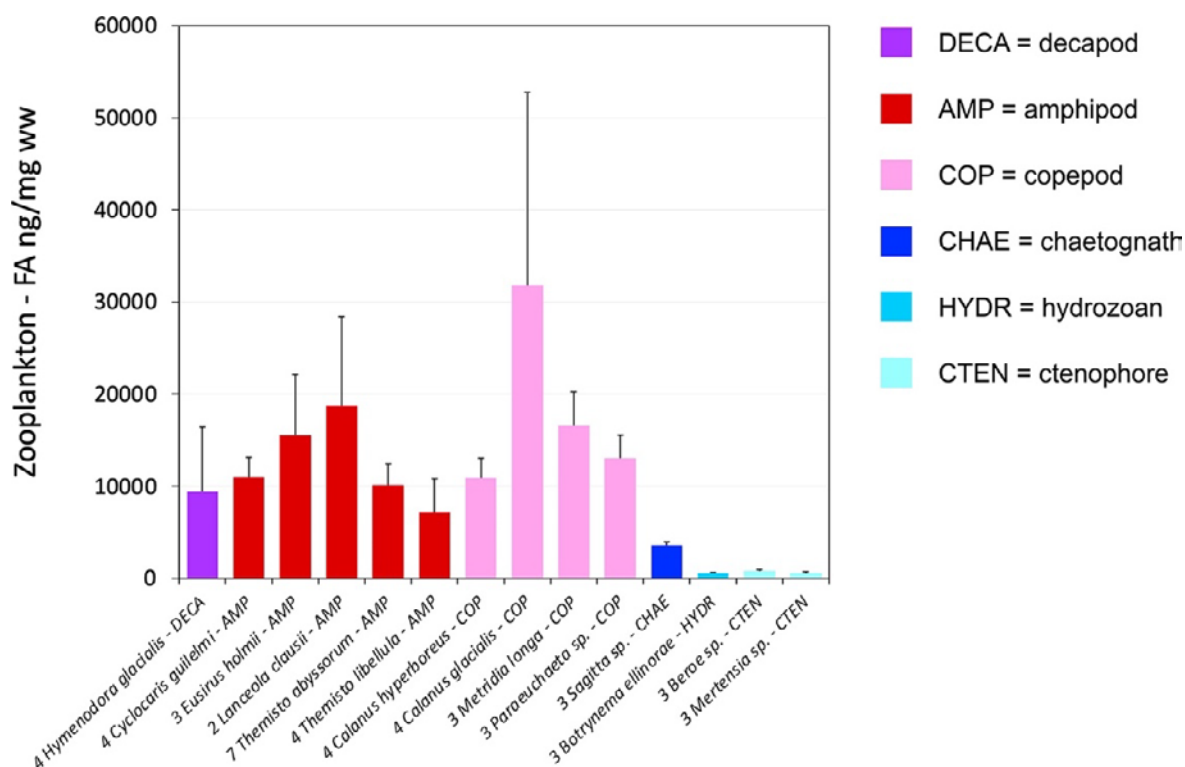
Figure 56: Fatty acid (FA) content in the 14 zooplankton taxa analysed. The number replicate zooplankton individuals of the same species is given before the taxon name on the X-axis. Different colours represent the different taxonomic groups as given after the taxon name on the X-axis. The error bars represent 1 standard deviation of the mean.
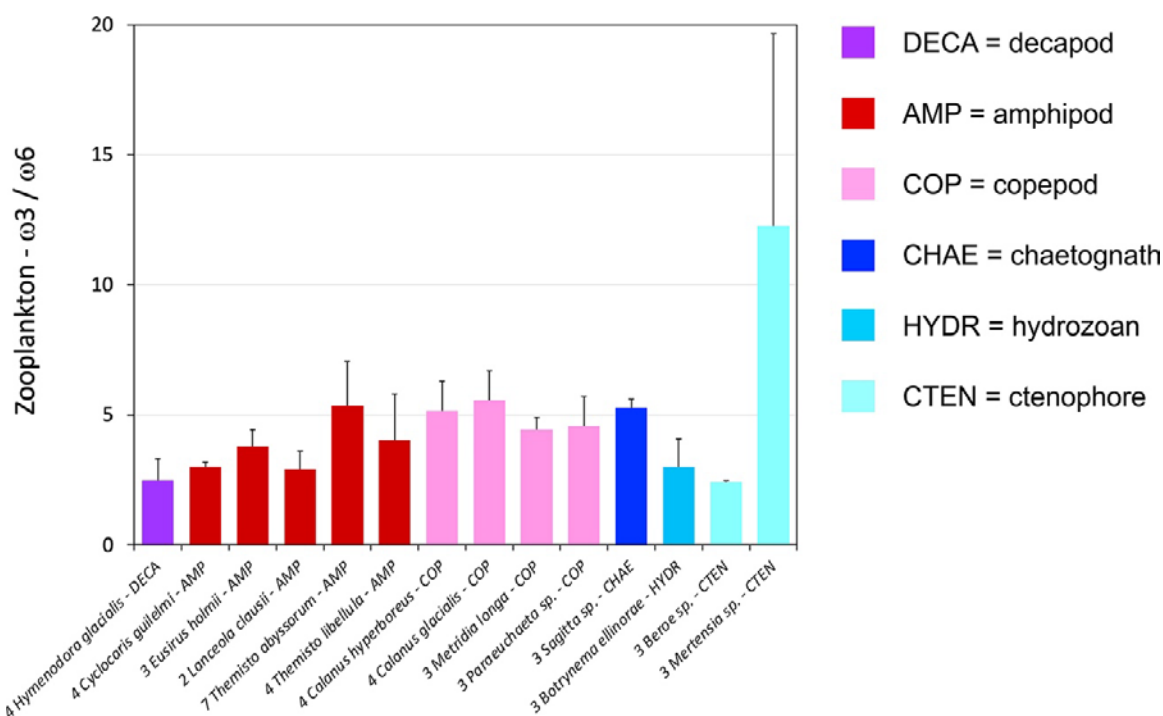
Figure 57: The ω3/ω6 ratio of the polyunsaturated fatty acids (PUFAs) in the 14 zooplankton taxa analysed. The number of replicate zooplankton individuals of the same species is given before the taxon name on the X-axis. Different colours represent the different taxonomic groups as given after the taxon name on the X-axis. The error bars represent 1 standard deviation of the mean.

The nutritional quality of the zooplankton, expressed as the ω3/ω6 ratio of the PUFAs, showed that the larger carnivorous ctenophore *Mertensia* sp. had the highest value (12.3), a similar level as fish (see WP5), albeit with a large standard deviation (**Figure 57**). Among the other taxa, *Themisto abyssorum*, all copepods and *Sagitta* sp. had high values (ca. 5). Given the high numbers of *Themisto abyssorum* found in the CAO on the FishCam videos[115] and the high abundance of this species in the fish stomachs during the MOSAiC expedition (see WP5), it can be concluded that this is a key species as fish prey in the deep scattering layer of the CAO with high nutritional quality expressed as ω3/ω6.

Altogether, 31 FAs were detected in the 50 zooplankton samples (**Table 31**). The dominant SAFAs were C14:0 (myristic acid) and C16:0 (hexadecanoic acid), the dominant MUFAs were C16:1n-7 (palmitoleic acid) and C18:1n-9 (elaidic acid), and the dominant PUFAs were FA 20:5n-3 (eicosapentaenoic acid, EPA) and 22:6n-3 (docosahexaenoic acid, DHA).

Higher EPA/DHA ratios of the amphipods *Cyclocaris guilelmi*, *Lanceola clausi*, and *Themisto libellula* suggest that these species had a more diatom-originated (potentially ice-associated) diet than the other taxa (**Table 29**, **Figure 58**). Standard deviations for FA composition are usually naturally high, but it is also a consequence of balancing the number of analysed species against the high analysis costs. In this case the goal of the study was to obtain an overview of the 14 main prey species in the fish food web.

The amphipod *Themisto libellula*, the copepod *Paraeuchaeta* sp. and the ctenophore *Mertensia* sp. had the highest PUFA/SAFA ratios (ca. 1.4), indicating more carnivorous feeding than the other taxa (**Table 31**, **Figure 58**). The omnivores *Themisto abyssorum* and *Metridia longa* had slightly lower values (ca. 1.2) and the mainly herbivorous *Calanus glacialis* had the lowest value (0.6). However, *Calanus hyperboreus* had relatively high values (1.2), and the carnivorous ctenophore *Beroe* sp. deviated from the pattern with a low average PUFA/SAFA ratio of 0.6. Standard deviations are high because the number of replicates per zooplankton taxon was low, and these results should be interpreted with care.

A Principal Components Analysis (PCA) was performed based on the relative shares of 5 fatty acid trophic markers (FATM) in the muscle tissue of fish. FATM are essential fatty acids which are synthesised by certain algae groups and prey taxa and cannot be synthesised by other animals in the food web. FATM are therefore preserved during trophic transfer from one trophic level to the next and assumed to be not metabolised or chemically modified. The following FATM[116] was used: 16:1n-7 and 20:5n-3 as indicators of diatom algae (Bacillariophyceae), 18:4n-3 and 22:6n-3 as indicators of flagellates (various taxa, including Dinophyceae), and 20:1n-9 as indicator of *Calanus* copepods.

The PCA biplot showed a gradient along PC1 from the diatom-associated FATM 16:1n-7 to the left and the flagellate-associated FATM 22:6n-3 and the second diatom-associated FATM 20:5n-3 to the right (**Figure 59**). Because 16:1n-7 is often enriched in diatoms from cold environments including sea ice[117], this gradient could be related to the relative importance of ice algae versus phytoplankton as a carbon source for the investigated zooplankton species. Accordingly, the ice-associated amphipod *Eusirus holmii*, and the partly ice-associated copepod *Calanus glacialis* grouped with the FATM 16:1n-7, as well as the copepod *Metridia longa* which can be abundant in the under-ice habitat[118].

[115] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

[116] Kohlbach D, et al. (2016) The importance of ice algae-produced carbon in the central Arctic Ocean ecosystem: Food web relationships revealed by lipid and stable isotope analyses. Limnology & Oceanography 61:2027-2044 [http://doi.org/10.1002/lno.10351]

[117] Kohlbach D, et al. (2016) The importance of ice algae-produced carbon in the central Arctic Ocean ecosystem: Food web relationships revealed by lipid and stable isotope analyses. Limnology & Oceanography 61:2027-2044 [http://doi.org/10.1002/lno.10351]

[118] David C, et al. (2015) Community structure of under-ice fauna in the Eurasian central Arctic Ocean in relation to environmental properties of sea-ice habitats. Marine Ecology Progress Series 522:15-32 [https://doi.org/10.3354/meps11156]
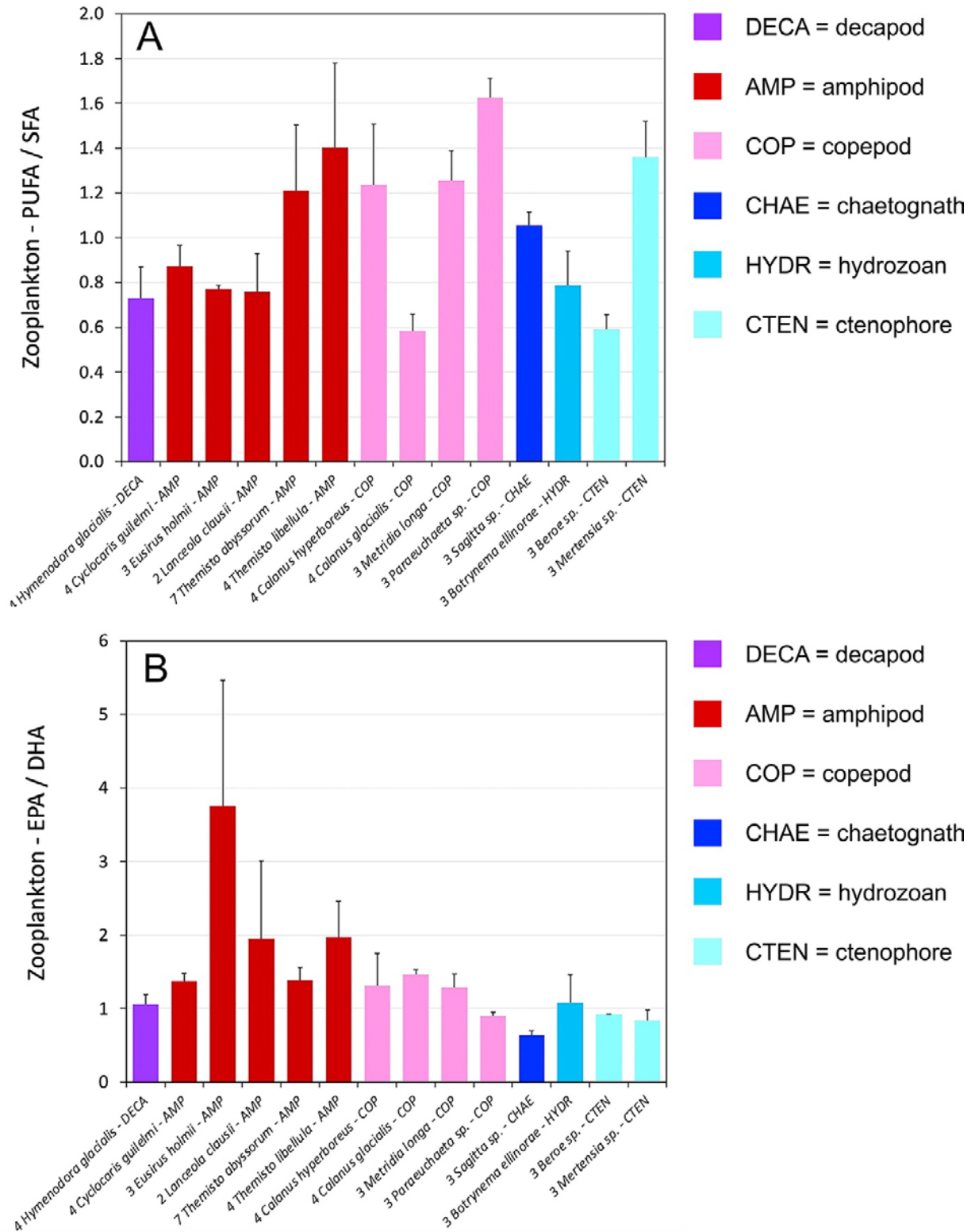
Figure 58: Composition of the fatty acid (FA) content in the 14 zooplankton taxa analysed. (A) the ratio of polyunsaturated and saturated fatty acids (PUFA/SAFA). (B) the ratio of the two polyunsaturated fatty acids 20:5n-3 (eicosapentaenoic acid, EPA) and 22:6n-3 (docosahexaenoic acid, DHA). The number of replicate zooplankton individuals of the same species is given before the taxon name on the X-axis. Different colours represent the different taxonomic groups as given after the taxon name on the X-axis. The error bars represent 1 standard deviation of the mean.

In contrast, the copepod *Calanus hyperboreus* and the pelagic amphipods *Themisto libellula* and *Themisto abyssorum* predominantly grouped with the flagellate-associated FATM 22:6n-3 and the diatom-associated FATM 20:5n-3, indicating a more phytoplankton-based diet. PC2 mainly correlated with the copepod-associated FATM 20:1n-9, indicating a copepod-based diet in the pelagic amphipods *Lanceola clausi* and *Cyclocaris guilelmi*, most of the jellyfish and several individuals of *Themisto* spp., the shrimp *Hymenodora glacialis* and the ice amphipod *Eusirus holmii*. In summary, these results confirm a high importance of sea-ice associated carbon in the prey species of polar cod, and to a lesser degree in the prey species of Atlantic fish, such as Atlantic cod and haddock (see WP5). Although they are not an important prey for Atlantic fish, copepods may be an important carbon source of their preferred amphipod prey, hence be indirectly relevant for large predatory fish.
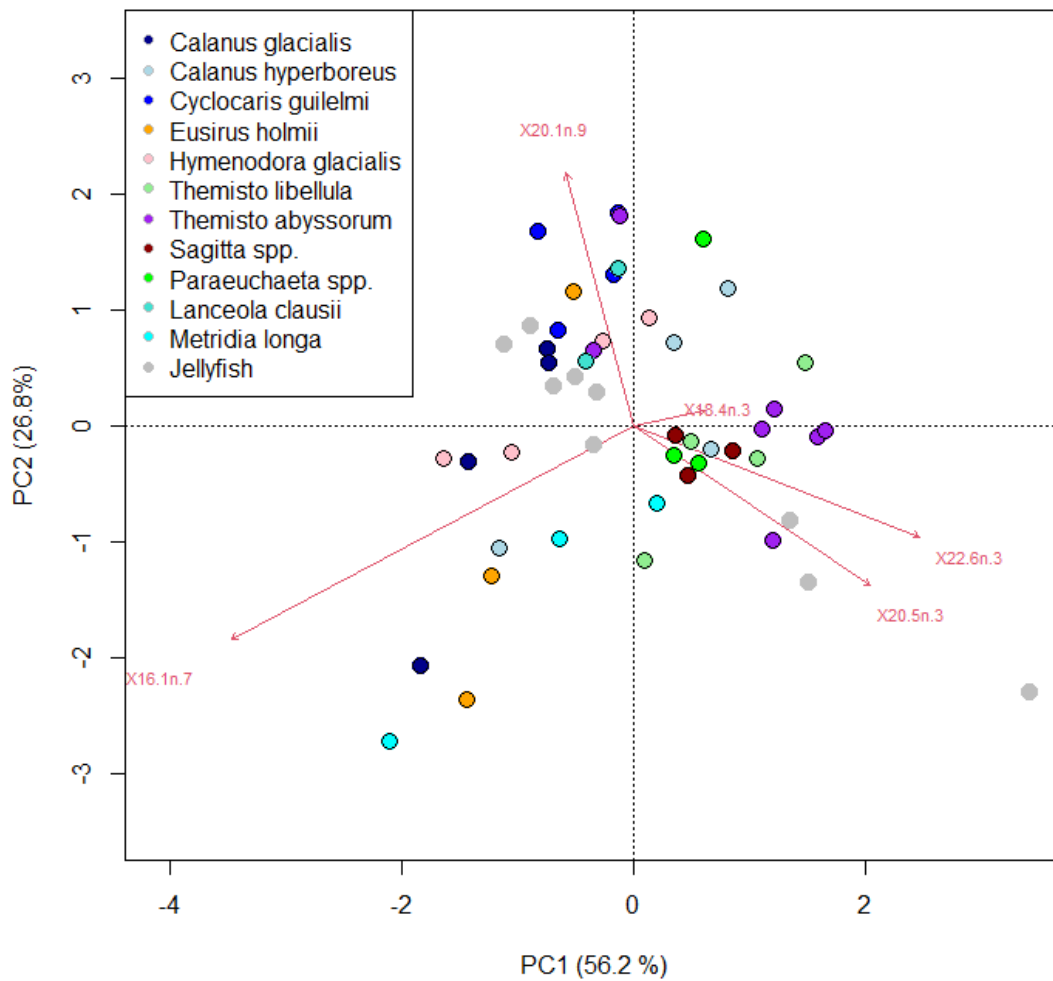
Figure 59: PCA biplot of the relative contribution of the 5 fatty acid trophic markers 16:1n-7, 20:5n-3, 18:4n-3, 22:6n-3 and 20:1n-9 in zooplankton species sampled during the SAS-Oden 2021 expedition.

Table 31: Results of the fatty acid analyses given as mean concentrations for each of the 31 fatty acids identified for each of the 14 zooplankton taxa. The data represent the mean ± 1 standard deviation of the mean.

	Cyclocaris guilelmi AMPHIPOD	*Eusirus holmii* AMPHIPOD	*Hymenodora glacialis* DECAPOD	*Lanceola clausii* AMPHIPOD	*Themisto abyssorum* AMPHIPOD	*Themisto libellula* AMPHIPOD	*Calanus hyperboreus* COPEPOD	*Calanus glacialis* COPEPOD	*Metridia longa* COPEPOD	*Paraeuchaeta* sp. COPEPOD	*Botrynema ellinorae* HYDROZOAN	*Sagitta* sp. CHAETOGNATH	*Beroe* sp. CTENOPHORE	*Mertensia* sp. CTENOPHORE
Nr of individuals	n=4	n=3	n=4	n=2	n=7	n=4	n=4	n=4	n=3	n=3	n=3	n=3	n=3	n=3
Saturated fatty acids (SAFA) in ng mg^{-1} wet weight														
12:0	60 ± 18	71 ± 29	139 ± 63	191 ± 144	150 ± 82	74 ± 47	134 ± 58	531 ± 285	95 ± 13	110 ± 13	5 ± 1	22 ± 7	7 ± 2	5 ± 1
3:0	9 ± 2	11 ± 2	12 ± 7	33 ± 31	19 ± 8	13 ± 14	15 ± 5	64 ± 36	14 ± 1	12 ± 2	1 ± 0	4 ± 0	1 ± 0	1 ± 0
14:0	411 ± 37	850 ± 369	488 ± 352	1200 ± 971	591 ± 136	415 ± 214	718 ± 159	3577 ± 2658	698 ± 127	590 ± 174	33 ± 4	179 ± 9	52 ± 5	47 ± 8
15:0	46 ± 10	66 ± 21	41 ± 27	110 ± 80	64 ± 20	40 ± 29	49 ± 7	216 ± 158	61 ± 1	50 ± 10	3 ± 0	18 ± 2	5 ± 1	3 ± 1
16:0	1376 ± 190	2812 ± 1389	1149 ± 753	3001 ± 2007	1522 ± 332	1031 ± 448	1406 ± 197	4655 ± 3450	2061 ± 381	1391 ± 163	84 ± 7	601 ± 46	137 ± 29	99 ± 21
17:0	16 ± 3	22 ± 6	14 ± 7	42 ± 29	29 ± 10	14 ± 10	16 ± 2	24 ± 4	34 ± 4	15 ± 1	1 ± 0	14 ± 2	2 ± 0	1 ± 0
18:0	107 ± 16	204 ± 90	122 ± 57	253 ± 173	128 ± 24	67 ± 20	97 ± 10	161 ± 32	136 ± 9	106 ± 9	12 ± 2	84 ± 7	15 ± 2	12 ± 3
20:0	3.5 ± 0.6	10.1 ± 0.2	4.3 ± 2.5	10.9 ± 7.8	5.5 ± 1.2	3.9 ± 1.9	4.7 ± 0.4	15.6 ± 13.7	5.9 ± 0.5	5.2 ± 1.0	0.3 ± 0.0	2.0 ± 0.1	0.5 ± 0.1	0.3 ± 0.1
21:0	0.4 ± 0.0	0.5 ± 0.2	0.5 ± 0.3	1.2 ± 0.9	0.6 ± 0.2	0.8 ± 0.7	0.5 ± 0.0	0.9 ± 0.4	0.7 ± 0.1	0.6 ± 0.1	0.0 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	0.0 ± 0.0
22:0	1.6 ± 1.1	11.7 ± 0.0	2.2 ± 1.2	8.9 ± 0.0	3.6 ± 1.0	1.6 ± 0.3	2.9 ± 0.3	5.4 ± 0.0	2.8 ± 2.5	2.8 ± 0.0	0.2 ± 0.0	1.2 ± 0.1	0.3 ± 0.0	0.1 ± 0.1
24:0	1.6 ± 0.1	9.2 ± 6.4	1.6 ± 1.0	4.6 ± 3.7	2.2 ± 0.7	0.9 ± 0.2	2.3 ± 0.3	7.3 ± 5.6	3.3 ± 0.5	2.5 ± 0.4	0.1 ± 0.0	0.9 ± 0.1	0.3 ± 0.1	0.1 ± 0.1
Monounsaturated fatty acids (MUFA) in ng mg^{-1} wet weight														
14:1n-5	26 ± 7	22 ± 8	28 ± 19	44 ± 18	23 ± 6	19 ± 13	74 ± 31	199 ± 159	37 ± 6	84 ± 38	2 ± 0	5 ± 2	1 ± 0	1 ± 0
16:1n-7	2107 ± 392	4299 ± 2295	2276 ± 2024	3589 ± 2086	1584 ± 190	1155 ± 324	2153 ± 724	9116 ± 7611	4720 ± 2450	2336 ± 246	132 ± 10	692 ± 106	166 ± 35	96 ± 26
17:1n-7	25 ± 7	22 ± 9	29 ± 24	48 ± 22	22 ± 6	14 ± 6	21 ± 10	63 ± 33	64 ± 3	45 ± 10	2 ± 0	8 ± 2	2 ± 1	1 ± 0
18:1n-9, trans	416 ± 110	588 ± 233	580 ± 481	854 ± 317	309 ± 55	233 ± 101	328 ± 72	838 ± 579	426 ± 60	343 ± 71	21 ± 3	148 ± 16	30 ± 8	13 ± 6
18:1n-9, cis	1696 ± 295	2000 ± 817	1974 ± 1748	2627 ± 1507	1241 ± 280	744 ± 238	992 ± 201	2055 ± 942	2700 ± 323	2052 ± 282	89 ± 11	380 ± 32	90 ± 19	45 ± 18
20:1n-11	1440 ± 393	886 ± 272	680 ± 429	1751 ± 359	752 ± 298	533 ± 350	969 ± 277	2534 ± 36	871 ± 67	1267 ± 813	53 ± 4	266 ± 76	71 ± 25	27 ± 15
22:1n-9	1368 ± 760	493 ± 136	463 ± 262	1428 ± 155	503 ± 154	277 ± 123	758 ± 189	2420 ± 2006	567 ± 50	836 ± 668	47 ± 17	177 ± 39	50 ± 11	23 ± 12
24:1n-9	63 ± 13	45 ± 19	48 ± 28	117 ± 43	58 ± 22	32 ± 8	97 ± 28	337 ± 273	157 ± 32	115 ± 42	5 ± 1	48 ± 17	6 ± 2	2 ± 1
Polyunsaturated fatty acids (PUFA) in ng mg^{-1} wet weight														
16:4n-3	42 ± 18	82 ± 84	17 ± 10	40 ± 35	76 ± 63	47 ± 67	37 ± 19	48 ± 10	43 ± 10	47 ± 11	1 ± 0	6 ± 1	2 ± 1	1 ± 0
18:3n-3 ALA	49 ± 6	102 ± 48	63 ± 38	137 ± 114	131 ± 81	101 ± 116	102 ± 44	142 ± 56	146 ± 11	166 ± 28	4 ± 0	41 ± 10	5 ± 1	3 ± 1
18:4n-3 SDA	87 ± 18	310 ± 147	99 ± 62	280 ± 225	395 ± 247	280 ± 331	614 ± 540	345 ± 81	435 ± 12	471 ± 160	9 ± 4	54 ± 4	11 ± 4	7 ± 2
20:3n-3	10 ± 4	17 ± 5	12 ± 7	39 ± 36	29 ± 17	14 ± 13	19 ± 8	15 ± 3	14 ± 2	16 ± 2	1 ± 0	5 ± 1	1 ± 0	1 ± 0
20:5n-3 EPA	657 ± 95	1573 ± 871	402 ± 200	1209 ± 227	1177 ± 458	961 ± 475	995 ± 361	2248 ± 1445	1425 ± 290	1098 ± 98	37 ± 19	279 ± 35	35 ± 5	90 ± 31
22:6n-3 DHA	477 ± 38	399 ± 71	370 ± 161	769 ± 537	871 ± 365	530 ± 371	773 ± 216	1509 ± 889	1108 ± 194	1222 ± 171	33 ± 9	436 ± 22	38 ± 6	108 ± 29
18:2n-6 LA	157 ± 23	240 ± 116	169 ± 125	276 ± 169	234 ± 109	148 ± 109	199 ± 107	302 ± 135	318 ± 27	283 ± 49	9 ± 1	60 ± 3	12 ± 2	6 ± 3
18:3n-6	19 ± 5	105 ± 73	12 ± 8	42 ± 27	26 ± 11	19 ± 7	27 ± 12	80 ± 65	67 ± 27	22 ± 1	1 ± 0	6 ± 1	1 ± 0	1 ± 0
20:2n-6	158 ± 34	176 ± 60	183 ± 126	445 ± 367	165 ± 42	335 ± 302	217 ± 32	244 ± 41	235 ± 39	315 ± 106	12 ± 1	70 ± 6	20 ± 3	10 ± 4
20:3n-6	15 ± 4	36 ± 21	10 ± 9	14 ± 8	8 ± 1	7 ± 3	12 ± 4	45 ± 39	17 ± 3	8 ± 1	0 ± 0	2 ± 0	1 ± 0	0 ± 0
20:4n-6 ARA	70 ± 34	54 ± 17	30 ± 16	77 ± 13	35 ± 7	25 ± 6	20 ± 7	48 ± 25	50 ± 14	17 ± 2	3 ± 1	9 ± 1	2 ± 1	1 ± 1
22:2n-6	20 ± 3	27 ± 7	28 ± 12	65 ± 61	19 ± 5	29 ± 29	22 ± 2	25 ± 3	27 ± 4	25 ± 4	2 ± 0	9 ± 2	2 ± 0	2 ± 0
Summary of concentrations in ng mg^{-1} wet weight														
sum FA	10934 ± 2198	15542 ± 6547	9446 ± 6944	18706 ± 9747	10173 ± 2270	7164 ± 3646	10875 ± 2140	31869 ± 20879	16539 ± 3748	13054 ± 2467	599 ± 49	3626 ± 342	767 ± 155	606 ± 112
sum SAFA	2032 ± 255	4067 ± 1895	1973 ± 1262	4855 ± 3448	2515 ± 575	1662 ± 773	2445 ± 395	9257 ± 6609	3111 ± 496	2284 ± 356	139 ± 13	926 ± 37	221 ± 39	168 ± 33
Sum MUFA	7141 ± 1816	8355 ± 3266	6078 ± 4993	10458 ± 4507	4491 ± 751	3008 ± 1105	5393 ± 1101	17563 ± 11569	9543 ± 2927	7078 ± 1712	349 ± 23	1725 ± 282	416 ± 100	209 ± 78
Sum PUFA	1762 ± 193	3120 ± 1429	1394 ± 738	3394 ± 1792	3167 ± 1315	2495 ± 1780	3037 ± 966	5049 ± 2723	3884 ± 483	3692 ± 402	111 ± 30	976 ± 45	130 ± 22	229 ± 54
Sum Omega3	1322 ± 135	2482 ± 1174	962 ± 464	2474 ± 1174	2680 ± 1191	1933 ± 1343	2540 ± 836	4306 ± 2463	3171 ± 454	3020 ± 452	83 ± 30	820 ± 40	92 ± 16	209 ± 56
Sum Omega6	440 ± 61	638 ± 265	432 ± 294	919 ± 618	487 ± 149	562 ± 449	497 ± 151	743 ± 266	713 ± 31	672 ± 71	28 ± 1	155 ± 10	38 ± 6	20 ± 8

4.6. Answers to the WP4 research questions

(1) How much fish biomass can be supported by the pelagic ecosystem of the CAO?

Assuming that the meso- and macrozooplankton biomasses can be added up since the nets capture different size fractions of a population, the total zooplankton biomass in the CAO did not exceed 40 g m^{-2} during the SAS-Oden expedition. The trophic transfer efficiency from one trophic level to another is in the order of 10% in most ecosystems, including the Arctic Ocean[119,120,121]. This implies that the zooplankton community could at maximum support 4 g fish m^{-2}, in theory thus up to 13.2 million tonnes of pelagic fish in the CAO (3.3 million km^2). The estimated mean pelagic fish biomass density in the water column between the surface and 600 m was between 20 and 230 kg km^{-2}, which equals 0.0003 - 0.004 g m^{-3} during the SAS-Oden expedition (see WP2). A tentative extrapolation to the entire CAO would imply that pelagic fish biomass was in the order of 66 000 to 759 000 tonnes based on the SAS-Oden survey. This value does not include the biomass of polar cod dwelling in the under-ice habitat, as under-ice prey biomass could not be quantified with the methods applied during the SAS-Oden expedition.

Although the order of magnitude of zooplankton biomass appears to have been sufficient to support the potential pelagic fish biomass during the SAS-Oden expedition, several aspects must be considered: Firstly, planktivorous fish such as myctophids and polar cod compete with other predators such as ctenophores, cnidarians, chaetognaths, carnivorous crustaceans and young life stages of the armhook squid *Gonatus fabricii*, reducing the zooplankton biomass available as fish diet. Secondly, this tentative calculation includes large uncertainties from both zooplankton biomass estimates and fish biomass estimates based on hydroacoustic data, and assumes homogenous distribution of zooplankton and fish in the entire CAO. To reduce this uncertainty, the species composition of pelagic fish must be known, and high-resolution spatial distribution models for fish and zooplankton must be applied to realistically estimate the distribution of fish and their zooplankton prey in the CAO.

(2) What is the quality of zooplankton species as food for fish?

The potential zooplankton prey species inhabiting the CAO are of similarly good quality in terms of carbon and nitrogen content and fatty acid composition as in other regions such as the Barents Sea and the North Atlantic where fish is commercially caught[122,123]. Thus, it can be concluded that not the food quality but the food quantity limits growth and reproduction in fish and other nektonic species such as squid.

(3) What are the trophic interactions between zooplankton prey species and fish? (collaboration with WP5)

Results from ^{13}C and ^{15}N stable isotope analyses of fish muscle tissue confirmed that polar cod were dependent on ice-associated zooplankton such as *Apherusa glacialis* and

[119] Lindeman RL (1942) The trophic-dynamic aspect of ecology. Ecology 23:399–417 [https://doi.org/10.2307/1930126]

[120] Flores H, et al. (2019) Sea-ice properties and nutrient concentration as drivers of the taxonomic and trophic structure of high-Arctic protist and metazoan communities. Polar Biology 42:1377-1395 [https://doi.org/10.1007/s00300-019-02526-z]

[121] Ehrlich J, et al. (2021) Sea-ice associated carbon flux in Arctic spring. Elementa: Science of the Anthropocene 9 [https://doi.org/10.1525/elementa.2020.00169]

[122] Bogstad B, et al. (2000) Who eats whom in the Barents Sea? NAMMCO Scientific Publications 2:98-119 [https://doi.org/10.7557/3.2975]

[123] Jørgensen LL, et al. (2020) Responding to global warming: New fisheries management measures in the Arctic. Progress in Oceanography 188:102423 [https://doi.org/10.1016/j.pocean.2020.102423]

Calanus glacialis. This impression was confirmed by the fatty acid pattern, in which diatom-associated fatty acids, indicative of ice algae, dominated over flagellate-dominated fatty acids, indicative of phytoplankton. In Atlantic cod, haddock, beaked redfish and ice cod, C and N stable isotope values clustered apart from polar cod and their zooplankton prey in the Arctic Ocean. This suggests that the time spent in the Arctic Ocean may not have been sufficient to transfer the isotopic signal of the Arctic pelagic food web to their muscle tissue, due to long turnover times. The relative marker fatty acid contributions showed that polar cod and black seasnail were predominantly associated with fatty acids indicative of diatoms. Haddock, Atlantic cod and beaked redfish caught in the Atlantic inflow region were associated with a stronger signal from pelagic flagellate fatty acids. Atlantic cod caught in the CAO showed a higher contribution of diatom-associated markers and low $\delta^{13}C$ values in relation to $\delta^{15}N$ levels, indicating a partial contribution of the cold-adapted food web of the CAO.

(4) How does zooplankton abundance estimated by net data relate to optical and hydroacoustic data? (collaboration with WP3)

Hydroacoustic methods mainly target the larger zooplankton group. Unfortunately, the deployments of the ring net that quantitatively captures this size class, were very limited, and the data do not allow for a comparison between the ring net and beam net results. For targeting mesozooplankton, the multinet Midi, UVP and LOKI were deployed. The results from the multinet yielded higher abundances than LOKI and UVP, and is considered the best method to estimate the mesozooplankton biomass available as prey for nekton. Surprisingly, mesozooplankton biomass data as calculated from the multinet samples and macrozooplankton biomass as calculated from acoustic data correlated significantly. This suggests that meso- and macrozooplankton abundances were closely correlated during the SAS-Oden cruise. Future studies are, however, needed to test to which extent mesozooplankton biomass can predict macrozooplankton biomass.

4.7. Relevance of the WP4 data for fish stock modelling

The here obtained zooplankton abundance, biomass and biochemical composition data are relevant for future fish stock modelling and assessment in the CAO as these parameters estimate the prey field for fish populations. Furthermore, the spatial correlation of zooplankton biomass and fish biomass emphasises the potential of zooplankton data as an indicator of fish distribution. The biomasses of the two zooplankton size classes sampled with nets, mesozooplankton and macrozooplankton, can directly be used in models[124] to predict trophic interactions in the CAO. The UVP data (WP3) support the net surveys and yield fine-scale zooplankton distribution patterns that allow identification of areas that could indicate rich feeding grounds for fish by high abundances of zooplankton.

With regard to stable isotope composition, C:N ratios and fatty acid composition, the here obtained data agree with previous studies from the CAO[125]. As the Arctic ecosystem is changing rapidly, it is important to monitor the long-term variability of the carbon flux, food web structure and nutritional value of zooplankton prey in the CAO in order to assess the future sustainability of potential fisheries. The collection of basic data on stable isotope composition and C:N ratios will be a valuable and cost-effective tool to monitor potential changes of the food web and the nutritional value of fish prey during

[124] Utne KR, et al. (2012) Estimating the consumption of *Calanus finmarchicus* by planktivorous fish in the Norwegian Sea using a fully coupled 3D model system. Marine Biology Research 8:527-547 [https://doi.org/10.1080/17451000.2011.642804]
[125] Kohlbach D, et al. (2016) The importance of ice algae-produced carbon in the central Arctic Ocean ecosystem: Food web relationships revealed by lipid and stable isotope analyses. Limnology & Oceanography 61:2027-2044 [http://doi.org/10.1002/lno.10351]

the coming years. These data may be complemented with fatty acid analysis and other trophic biomarkers (e.g., stable isotope composition of essential fatty acids and essential amino acids), when sufficient resources are available.

4.8. Recommendations from WP4 for the JPSRM of the CAOFA

Deploy a Surface and Under-Ice Trawl (SUIT). This trawl was designed for under-ice fish (juvenile polar cod) and ice invertebrates[126].

Design a standard "JPSRM macrozooplankton light-trap line". In the SC07 project it was shown that macrozooplankton species, especially amphipods, are major food items for the mesopelagic fish in the CAO, but these larger animals are rarely caught in multinet samples (targeting mesozooplankton). To estimate access to food for mesopelagic fish, it would be highly recommended to develop a standard light-trap line to be deployed at standard depths in different areas of the CAO for a specific period of time. The macrozooplankton is attracted fast by light (scale of minutes) and deployment time could be one to several hours, but should be tested, The line should cover at least 0-800 m of depth and have weights (to keep the line vertical), a standard-type zooplankton light trap (commercial or especially designed) every 100 m, and a depth/temperature sensor at 400 and 800 m (to check that the line is kept vertical). Amphipods were attracted by red light as shown by the FishCam on the CTD during the SAS-Oden expedition, but other colours of light could be tested before defining the standard line that will be used within the JPSRM all over the CAO. This will also provide animals for C:N ratios, stable isotopes and fatty acids analyses from specific depths with higher precision than a MIK net.

Design a standard "JPSRM macrozooplankton sampling net" This should be a closable net so that it is possible to choose a specific depth. During the SAS-Oden expedition a MIK net with a standard diameter of 2 m was used (not closable). This is was large enough to sample macrozooplankton in a representative manner. The usual ring nets (e.g., Nansen net, WP2) are too small, and bigger ones (like the MIK net) do not have a closing mechanism. It should be possible to use a Nansen release mechanism on a large ring net, such as the WP3. Always use standard depths that cover the whole DSL and a bit below, e.g., 1000-500 m, 500-200 m, 200-0 m (to be decided by the JPSRM).

Use a standard "JPSRM mesozooplankton multinet". For example, the much used "Midi". Decide upon standard depth strata: commonly use in the CAO are the depth intervals 2000-1000-500-200-100-0 m (5 nets). Decide upon standard mesh size – commonly used in the CAO is 150 μm. This net targets mesozooplankton, i.e., the prey of smaller fish such as juvenile polar cod and myctophids.

Use standardised image-based measurements for zooplankton. The SC07 project results highlight the potential of computer-aided taxonomic analysis using "ZooScan" in combination with deep-learning software as a time-efficient method to quantify zooplankton species composition, abundance and biomass in the context of a monitoring programme such as the JPSRM. The by the SC07 project established training set on "EcoTaxa" will speed up future development of image analyses of macro- and mesozooplankton of the CAO.

---

[126] Van Franeker JA, et al. (2012) The Surface and Under-Ice Trawl (SUIT). Technical Report
https://www.researchgate.net/publication/297794282_The_Surface_and_Under_Ice_Trawl_SUIT#fullTextFileContent

5. FISH SAMPLES (WP5)

5.1. Research questions addressed by WP5

(1) Which species of fish occur in the CAO?
(2) What is the origin of the fish in the CAO?
(3) What is the diet, the condition and the nutritional quality of the fish?
(4) What are the trophic interactions between zooplankton prey species and fish?
   (collaboration with WP4)

5.2. Data produced by WP5

Data file "EFICA_DATA_SC07-WP5"

For details of the Device Operations (date, time, geographical position, station depth), see files "MOSAiC_Device_operations" and "SAS-Oden_2021_Logbook"[127].

5.3. Human resources of WP5 and main responsibilities

Hauke Flores (AWI) coordination of the fish sample analyses, fish dissection, stomach analyses, stable isotopes ($\delta^{13}C$, $\delta^{15}N$) and C:N ratio analyses and calculations; Fokje Schaafsma (WMR) stomach analyses; Kim Vane (AWI) otolith analyses, Filip Volckaert, Sarah Maes, Marie Verheye (KUL) population genetic analyses, Martina Vortkamp (AWI) stable isotope and C:N ratio analyses; Pauline Snoeijs-Leijonmalm (SU) fatty acids analyses and calculations

5.4. Methods used by WP5

Condition index, hepatosomatic index, gonado-somatic index

Total length (TL), standard length (SL) and wet mass (WM) of each fish were measured on board following the procedure described in SC-03 (MOSAiC). To investigate the condition of the eight Atlantic cod, 38 haddock, four beaked redfish and one black seasnail caught during MOSAiC Leg 3 and Leg 4 and the four polar cod caught during SAS-Oden, a basic dissection was performed. During this dissection, samples were also retrieved for population genetic, stomach content, otolith and stable isotope analyses. Nineteen of these fish had already been dissected during the MOSAiC expedition, and three out of four polar cod were dissected on board of Oden. Thirty-six fish not dissected during the MOSAiC expedition and one polar cod from the SAS-Oden expedition were dissected in the laboratory at AWI.

The wet mass (WM) of each fish was measured to the nearest 0.1 g before the dissections started. A piece of the pelvic fin was collected for genetic analyses and preserved on molecular grade ethanol. Then, the fish body was cut open to remove the inner organs (gonads, stomach, intestine, liver). The sex was determined according to gonad morphology. After the removal of all organs, the eviscerated mass (EM) was measured to the nearest 0.1 g. The stomach and intestine were preserved in 96% molecular grade ethanol. The liver was weighed to the nearest 0.1 g, as well as the gonads. The gonads were preserved in 4% formaldehyde solution. The otoliths were removed and stored dry in plastic vials.

---

[127] Bolin Centre Database [https://bolin.su.se/data/]

For each fish, Fulton's condition index (K) was calculated as:

$$K = 100 * WM/TL^3 \qquad \text{(Equation 10)}$$

Where WM is the total fish wet mass and TL is the fish total length

The hepatosomatic index (HSI) was calculated as:

$$HSI = 100 * LM/WM \qquad \text{(Equation 11)}$$

where WM is the total fish wet mass and LM is the fish liver mass

The gonadosomatic index (GSI) was calculated as:

$$GSI = 100 * GM/WM \qquad \text{(Equation 12)}$$

where WM is the total fish wet mass and GM is the fish gonad mass

Stomach contents

The stomachs of three haddocks and one beaked redfish were damaged and could not be analysed. The four polar cod sampled during the SAS-Oden expedition were not analysed because they were sampled with 24-hour exposed baited traps and their stomach contents would reflect the bait and not their natural food items. In total, the stomach contents of eight Atlantic cod (*Gadus morhua*), 35 haddock (*Melanogrammus aeglefinus*), three beaked redfish (*Sebastes mentella*) and one black seasnail (*Paraliparis bathybius*) were analysed according to the procedure previously described[128]. Prior to microscopic analysis, the ethanol-preserved fish stomachs were opened with a scalpel or scissors, and the content was rinsed through a 30-µm gauze. Then, the sample was rinsed into a clean glass Petri dish or directly into a Bogorov counting chamber, depending on the size of the stomach and contents. Recognizable food items were identified to the lowest taxon possible and enumerated using a high-quality stereo microscope equipped with a digital camera and an image analysis software.

Images of the prey items were taken and processed with image analysis software. WMR used a Zeiss V8 stereo microscope with an AxioCamHRc camera and the "Zen Imaging" software. AWI used a Leica M205C stereo microscope with a Leica MC170 HD camera and the software package "Leica Application Suite" version 4.12. A minimum number of individuals (MNI) of each prey item in each fish stomach was estimated for those prey taxa where it was possible, using both whole animals and identifiable body parts. For example, numbers of amphipods were estimated by adding the numbers of complete animals and the numbers of identifiable amphipod telsons. For fish, separate body parts, such as eyes, vertebrae or scales, were assumed to belong to a single fish unless there was evidence indicating otherwise, for example, when there were more than two eyes. Minimum numbers of chaetognaths were estimated based on the number of complete heads. After identification and enumeration, stomach contents were stored in molecular grade ethanol, to allow for potential future genomic-based diet analysis.

The average MNI was calculated for each prey taxon per fish species. In addition, the frequency of occurrence (FO) of each prey category in the stomachs of each of the fish individuals was calculated by dividing the number of fish with that particular prey category in the stomach by the total number of fish and then multiplying by 100 to obtain percentage values.

Stable isotopes and C:N ratios

---

[128] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

During both the MOSAiC and the SAS-Oden expeditions, muscle tissue of all fish caught was collected to perform C:N and stable isotope analysis of carbon (C) and nitrogen (N). A piece of muscle tissue from each fish was transferred to a pre-weighed tin cap and dried for 24 hours in a freeze dryer and ground to powder. After drying, the tin caps containing the tissue samples were weighed to determine the dry mass (DM) of the sample, using a calibrated microscale. The tin caps were then closed, sent to the Institute Littoral Environnement et Sociétés (LIENSS) at La Rochelle, France, and ratios were calculated (see WP4).

Fatty acid analyses

The fatty acid (FA) content and composition was used to assess the condition of the fish and as a biomarker for food-web interactions between zooplankton and fish (see WP4). Fatty acids in fish muscle were quantified in 30 fish individuals, eight from the CAO (one *Paraliparis bathybius*, one *Arctogadus glacialis*, three *Boreogadus saida*, three *Gadus morhua*), and 22 from the Atlantic inflow region near the Yermak Plateau and in Fram Strait (nine *Gadus morhua*, 12 *Melanogrammus aeglefinus*, one *Sebastes mentella*). All fish were sampled during the MOSAiC expedition except for the *Boreogadus saida*, which were sampled during the SAS-Oden expedition. This includes all fish from the CAO, but not all fish from the Atlantic inflow region, because not all fish could be dissected on board (FA samples must be frozen at -80 °C immediately after sampling). Fatty acids (FA) in fish liver were quantified in 14 fish individuals, three from the CAO (one *Arctogadus glacialis*, two *Gadus morhua*), and 11 from the Atlantic inflow region (seven *Gadus morhua*, three *Melanogrammus aeglefinus*, one *Sebastes mentella*). All fish liver samples were taken during the MOSAiC expedition.

The fish samples were taken out of the -80 °C freezer, and two sub-samples of ca. 20 mg of fresh weight were taken from each sample. Each sub-sample was placed into a separate Eppendorf tube and freeze-dried before extraction. The chemical analyses were performed with gas chromatography – mass spectrometry (GC-MS) at the Swedish Metabolomics Centre (SMC) in Umeå as described in WP4. After the analyses, the mean value of two sub-samples per sample was used to calculate means per fish taxon / sampling area.

Age reading

Before age reading, a very small (ca. 30 µg) of powder sample was taken from the outer layer of each otolith with a hand drill for $\delta^{13}C$ and $\delta^{18}O$ analysis (**Figure 75**). Otolith sections were made by breaking the otoliths through the core and embedding both pieces in VISCOVOSS GTS polyester resin. Sections of 0.5 mm were cut with a double-bladed diamond saw at the end of each otolith break line. Two otolith sections were recovered from each otolith, where feasible, and glued to a glass slide. High-resolution photographs of the otoliths were taken under a stereo microscope (Leica M205) connected to a digital imaging system based on the software Leica Application Suite, version 4.12. Based on the sequence of translucent and opaque increments, the number of years each fish had lived before it was sampled was determined. Winter increments are usually more translucent than summer increments. A year of life was counted as complete when both the opaque summer increment and the translucent winter increment were visible. The ages of the Atlantic cod, haddock, ice cod and polar cod were estimated with this method. The otoliths of redfish and black seasnail have a more complex inner structure that impedes age reading using this method.

Temperature reconstructions

Based on $\delta^{18}O$ analysis of otolith increments of gadoid fish, temperature reconstructions of the four gadoid fish species (Atlantic cod, polar cod, haddock and ice cod) of the specimens so far not analysed were performed. Results from the concurrent analysis of $\delta^{13}C$ values were used to investigate the potential of these data to calculate field metabolic rates.

Samples from summer and winter growth increments were micro-milled with a 0.8 mm diamond encrusted drill bit to a depth of 100 µm. These low-volume otolith increment samples were directly sampled into glass vials designed for isotopic measurements. Otolith powder from each increment was analysed for $\delta^{18}O$ and $\delta^{13}C$ values on a ThermoFisher 253plus gas isotope mass spectrometer connected to a Kiel IV automated carbonate preparation device at MARUM, Center for Marine Environmental Sciences, University Bremen (Germany)[129]. Data were reported in the usual delta notation versus V-PDB. The instrument was calibrated against the NBS 19 calcite standard. Over the measurement period, the standard deviations of the house standard were 0.02 ‰ for $\delta^{13}C$ and 0.07 ‰ for $\delta^{18}O$. For the Atlantic gadoids Atlantic cod and haddock, otolith $\delta^{18}O$ values were used to reconstruct ambient temperatures (in °C) with an equation established for Atlantic cod in controlled experimental settings[130,131]:

$$T = (\delta C - \delta W) * 0.2^{-1} + 19.5 \hspace{3cm} \text{(Equation 13)}$$

where δC constitutes the otolith $\delta^{18}O$ value, and δW represents the mean Atlantic water $\delta^{18}O$ value of 0.2345 ‰ derived from the Global Seawater Oxygen-18 Database ver. 1.22[132].

Polar cod dwell often in surface waters near the sea-ice underside, which has much lower $\delta^{18}O$ values compared to Atlantic water. An initial data exploration of $\delta^{18}O$ values from polar cod otolith increment samples indicated a wide spread from values below 0 to 3.5, indicating that this wide value range reflected different mixtures of under-ice water, CAO surface water and Atlantic water. The outer increments of the polar cod caught during MOSAiC and the SAS-Oden expedition were used as a basis to estimate δW in each otolith increment. Otolith increment $\delta^{18}O$ values <0 were removed from the data set to exclude the effect of potentially erroneous measurements. δW was modelled as a function of δC using the following reference points:

(1) Surface water in which the polar cod were caught, assuming a mean surface water temperature of -1.3 °C[5] (δW = -2.8) versus the mean δC value of outer edge samples from the polar cod (δC = -1.4),
(2) minimum of CAO surface water (δW = -2.9[5]) versus the difference of mean and standard deviation of δC (δC = 1.0), and
(3) Atlantic water (δW = 1.234[5]) versus the maximum δC value of Atlantic cod dwelling in the CAO ($\delta^{18}O$ = 3.7),

The resulting linear relationship

$$\delta W = 1.21 * \delta C - 4.27 \; (p<0.01; R^2=0.99) \hspace{2cm} \text{(Equation 14)}$$

was used to estimate δW in each polar cod otolith increment.

[129] Foster LC, et al. (2008) Effects of micromilling on $\delta^{18}O$ in biogenic aragonite. Geochemistry, Geophysics, Geosystems 9, 1–6 [https://doi.org/10.1029/2007GC001911]
[130] Kim ST, O'Neil JR (1997) Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates. Geochimica et Cosmochimica Acta 61:3461–3475 [https://doi.org/10.1016/S0016-7037(97)00169-5]
[131] Høie H, et al. (2004) Temperature-dependent fractionation of stable oxygen isotopes in otoliths of juvenile cod (Gadus morhua L.). ICES Journal of Marine Science 61:243–251 [https://doi.org/10.1016/j.icesjms.2003.11.006]
[132] Schmidt, et al. (2009) Global Seawater Oxygen-18 Database - v1.22 (2009). [https://data.giss.nasa.gov/o18data/]

To visualise experienced temperatures during the life history of each fish, the reconstructed temperature was plotted against estimated age based on otolith readings for all fish from which measurements of at least three consecutive increments were available. To identify likely dispersal routes of fish entering the Arctic Ocean, life-history temperature trajectories were used in combination with known temperature ranges of Arctic and North Atlantic water masses as previously described[5] and geographic distribution ranges of source populations derived from population genetic analyses.

Population genetics

Fourteen samples of Atlantic cod (*Gadus morhua*) were collected at six stations during the PS122 MOSAiC expedition of RV *Polarstern* in 2019/2020. 38 haddock (*Melanogrammus aeglefinus*) individuals were collected during the same expedition at seven stations. A total of 21 samples of polar cod (*Boreogadus saida*) were collected during PS122 at seven stations of the CAO and during the Synoptic Arctic Survey (SAS) expedition of RV *Oden* in 2021 at four stations in the CAO. Adult and juvenile fish were identified morphologically by experts onboard.

DNA was extracted from fin clips of Atlantic cod, haddock and polar cod using the NucleoSpin® tissue kit. Atlantic cod were genotyped at the *Pantophysin I* (*PanI*) locus through Sanger sequencing according to Johansen et al. (2018). SNP genotypes of haddock were called using the MassARRAY iPLEX Platform (Agena Bioscience) according to Berg et al. (2021). Genotyping-by-sequencing (GBS) libraries were prepared for polar cod using the *PstI* restriction enzyme and sequenced on the Illumina Novaseq platform 6000 of the Genomics core of KU Leuven. The reference-based *Stacks* v.2.5 pipeline was then used to process GBS data into single-nucleotide polymorphisms (SNP) genotypes. First, raw sequence reads were demultiplexed and cleaned using the process_radtags module. *Bowtie2* v.2.5.0 was used for aligning sequencing reads to the Atlantic cod reference genome for comparison. Only SNPS genotyped in at least 70% of the individuals within each sampling site, a maximum heterozygosity of 0.70 and an overall minor allele frequency (MAF) of >0.05 were called using the population module. Further SNP filtering was done using *VCFtools* v.0.1.12 and R packages *poppr* v.2.9.3 and *hierfstat* v.0.5-11 (removal of monomorphic loci, loci with more than 20% missing data and/or heterozygosity >0.6 and individuals with more than 30% missing data). Departures from Hardy-Weinberg equilibrium were tested locus by locus in each population with R package *Pegas* v.0.11.12.

To allocate the haddock and polar cod sampled by the EFICA Consortium to potential source populations, the sequencing data were compared with larger data sets from publicly available databases (haddock: Berg et al. 2021) and from sample collections of AWI and KUL (polar cod).

Observed and expected heterozygosity (H_o and H_e) and inbreeding coefficients (F_{IS}) were calculated for each putative population of Atlantic cod, haddock and polar cod using the function 'basic.stats' [R package *hierfstat*]. Pairwise F_{ST} values and respective p-values were assessed with the function *gl.fst.pop* [R package *DartR* v.2.0.4.] after 1 000 permutations and clustered with a classic multidimensional scaling (MDS) analysis.

A Principal Components Analysis (PCA) was performed using the function *dudi.pca* [R package *ade4* v.1.7.19]. For haddock, the Discriminant Analysis of Principal Components (DAPC) was used to visualise a priori defined geographical groupings, i.e., the three clusters of Berg et al. 2021 plus one cluster for the samples collected by the EFICA Consortium. For polar cod, the DAPC was performed without defining populations *a priori*. Instead, successive K-means with an increasing number of clusters were performed on the PCA-transformed data, and the optimal number of clusters was defined using the Bayesian Information Criterion (BIC). This k-means procedure was performed using the R

function *find.clusters*, while the DAPC was computed using the R function *dapc* [R package *adegenet* v.2.1.8]. The number of PCs to retain was determined by comparing mean a-scores of the populations for each number of retained PCs, using the R function *optim.a.score* [R package *adegenet* v.2.1.8].

5.5. Results and discussion of WP5

Fish parameters

Part of the results of this chapter have already been published[133], and the publication is part of this report. Within the WP07 project 55 fish sampled during MOSAiC Legs 2-4 and the SAS-Oden expedition were analysed. This excludes the six Atlantic cod (three from Leg 1 and three from Leg 3 of MOSAiC) and the single ice cod (MOSAiC Leg 1) for which For a complete overview, the latter seven fish were also included in the SC07 data set. For a number of the small ice-associated juvenile *Boreogadus saida* dissected on board, the exact weights of internal organs could not be measured because gonad- and liver tissues were sometimes damaged during sampling or too small for measurements for the scales available in the ship. **Table 32** presents a summary of basic data for all pelagic fish sampled by during the MOSAiC expedition, as well for the four sympagic polar cod caught during the SAS-Oden expedition.

Atlantic cod: The 14 Atlantic cod caught during MOSAiC Leg 1 inside the CAO (3 individuals), Leg 3 near the Yermak Plateau (three individuals) and Leg 4 in Fram Strait (8 individuals) had a mean total length of 53.6 cm (range 32.2-73.0 cm) and a mean weight of 1.33 kg. Their mean condition index K was 0.7, their mean GSI (n=5) was 0.6, and their mean HSI (n=6) was 2.3 (**Table 32**). There were no significant differences in total length or condition index K between the MOSAiC expedition Legs 1 and 3 (**Figure 60 A, B**).

Haddock: The 38 haddock caught during MOSAiC Leg 3 near the Yermak Plateau (6 individuals) and Leg 4 in the Fram Strait (32 individuals) had a mean total length of 51.0 cm (range 30.5-83.0 cm) and mean weight of 1.51 kg. Their mean condition index K was 0.9, their mean GSI (n=21) was 0.8, and their mean HSI (n=34) was 2.3 (**Table 32**). There were no significant differences in total length or condition index K between the two MOSAiC expedition legs (**Figure 60 C, D**).

Beaked redfish: The four beaked redfish caught during MOSAiC Leg 3 near the Yermak Plateau (2 individuals) and Leg 4 in Fram Strait (2 individuals) had a mean total length of 40.6 cm (range 31.0-54.2 cm) and mean weight 0.69 kg. Their mean condition index K was 1.0, their mean GSI (n=3) was 0.6, and their mean HSI (N=3) was 1.5 (**Table 32**).

Polar cod: The four polar cod (*Boreogadus saida*) caught in the CAO during the SAS-Oden expedition had a mean total length of 13.2 cm (range 10.7-16.4 cm). The 18 polar cod caught during the MOSAiC expedition in the CAO had a mean total length of 11.0 cm (F. Schaafsma, unpublished data). The polar cod from the SAS-Oden expedition were slightly larger than those caught during the MOSAiC expedition, but the number of individuals is far too low for a statistical comparison. Only one polar cod was dissected on board the SAS-Oden expedition. This fish was a gravid female with length 16.4 cm, weight 15.9 g, condition index 0.8 and GSI 12.6 (**Table 32**).

---

[133] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

Ice cod: The single ice cod (*Arctogadus glacialis*), that possibly is a Walleye pollock (*Gadus chalcogrammus*) and needs further genetic analysis, caught in the CAO during the MOSAiC expedition (Leg 1) had total length 33.2 cm, weight 210 g and condition index 0.6 (**Table 32**).

Black seasnail: The single black seasnail (*Paraliparis bathybius*) caught in the CAO during the MOSAiC expedition (Leg 2) had a total length 24.5 cm, weight 130 g and condition index 0.9 (**Table 32**).

Table 32: Summary of the dissection results for all fish caught during the MOSAiC and SAS-Oden expeditions. The data are shown as mean ± 1 standard deviation. K = condition index, GSI = gonadosomatic index, HSI = hepatosomatic index.

Species	N	Total length (cm)	Length range (cm)	Wet mass (g)	Gonad mass (g)	Liver mass (g)	K (index)	GSI (index)	HSI (index)
Atlantic cod	14	54 ± 13	32-73	1328 ± 892	8 ± 8	29 ± 25	0.7 ± 0.1	0.6 ± 0.3	2.3 ± 1.2
Haddock	38	51 ± 17	31-83	1506 ± 1341	13 ± 15	27 ± 27	0.9 ± 0.1	0.8 ± 0.7	2.3 ± 2.6
Beaked redfish	4	41 ± 10	31-54	690 ± 295	6 ± 5	12 ± 10	1.0 ± 0.3	0.7 ± 0.3	1.5 ± 0.6
Black seasnail	1	25	-	130	4.9	5.3	0.9	3.8	4.1
Ice cod	1	33	-	210	-	-	-	-	-
Polar cod (SAS-Oden)	4	13 ± 2	11-16	15.9	-	-	0.8	12.6	-
Total number of samples	80	78	78	75	34	44	75	31	44

Figure 60: Comparisons of total fish length and condition index K between different legs of the MOSAiC expedition for (A,B) Atlantic cod, and (C,D) haddock. Data are shown as standard box plots. The median value is indicated as a horizontal bar. The upper and lower margins of the "box" indicate the 25 and 75 percentiles, respectively, and the error bar shows 1.96 standard deviations. Outliers (extreme values) are not shown (see Table 32).

Fish diet

Within the WP07 project the stomach contents of 48 pelagic fish sampled during MOSAiC Legs 2-4 and the SAS-Oden expedition were analysed. The results for six Atlantic cod (MOSAiC Legs 1 and 3) and the single ice cod (MOSAiC Leg 1) have already been published. However, as these fish are also part of the SC07 data set, the previously published data are included here as well.

Atlantic cod: In terms of numbers, the diet of the Atlantic cod was dominated by
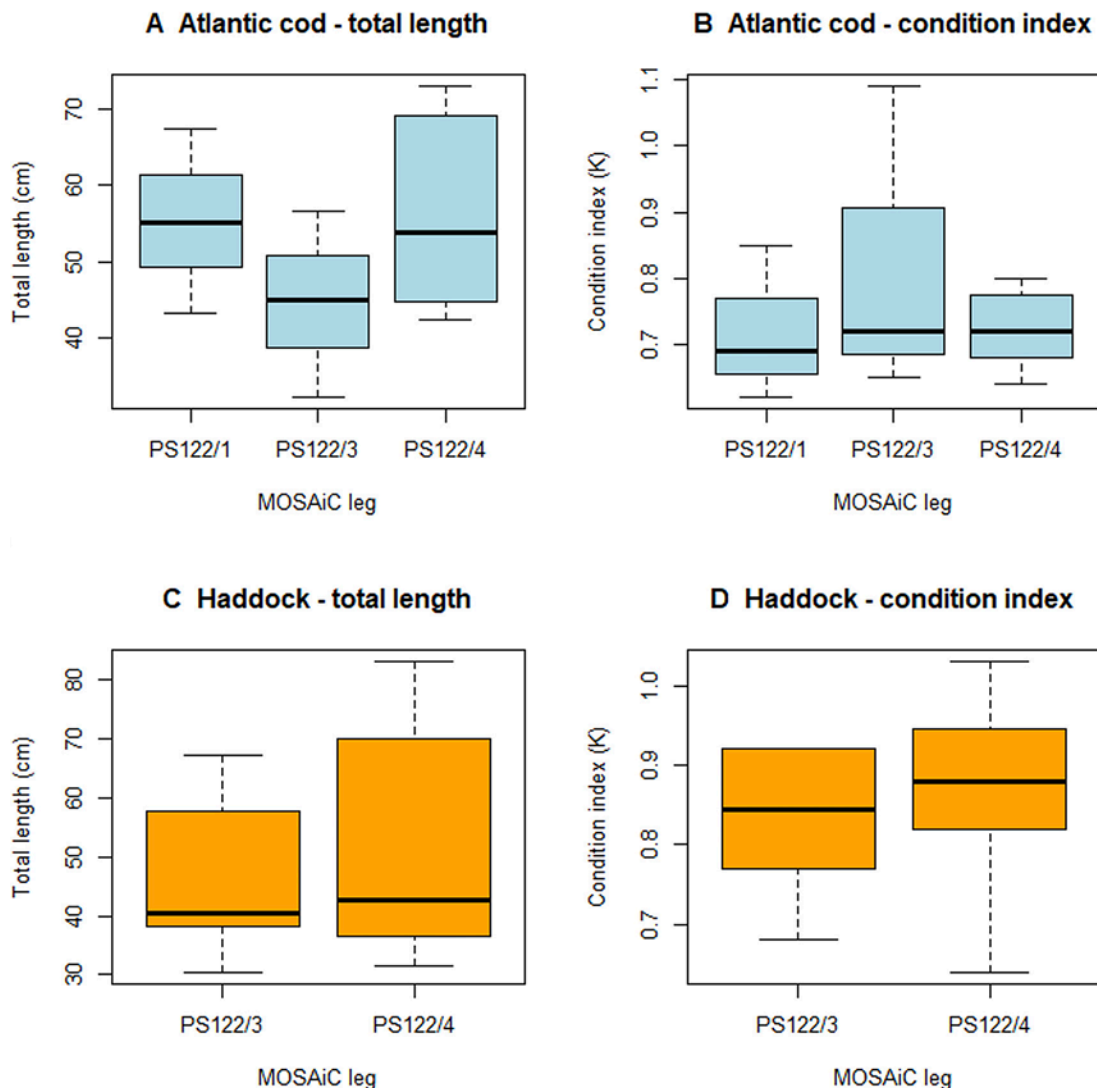
amphipods, followed by chaetognaths and fish (**Figure 61**, **Table 33**). 57% of the stomachs contained amphipods, 21% contained fish and 14% contained chaetognaths. Species-wise, the identifiable amphipods included *Cyclocaris guilelmi, Themisto libellula*, *Themisto abyssorum* and *Eusirus holmii*, and the chaetognaths were mainly *Sagitta maxima* (**Table 34**). 79% of the stomachs contained parasites, most of which were

trematode parasites with an average of 10.7 per stomach. Three single nematode parasites were found in three of the stomachs. Two fish stomachs contained the ectoparasitic copepod *Sphyrion lumpi*. One stomach from the Yermak Plateau contained a stone.

Haddock: In terms of numbers, the diet of the haddock was dominated by chaetognaths, followed by amphipods, ostracods, sea angels (Gymnosomata) and fish (**Figure 61**, **Table 33**). Chaetognaths occurred in 56% of the stomachs (**Table 34**). Identifiable amphipods from haddock stomachs included *Themisto abyssorum, Eusirus holmii* and *Themisto libellula*. Based on body parts, the presence of the amphipod species *Hyperoche medusarum, Hyperia medusarum* and/or *Hyperia galba* is suspected, but not enough features were visible to confirm identification.

Measurements performed on amphipods from haddock stomachs identified as *T. abyssorum* showed that the size of this species ranged from 12.8 to 18.0 mm. The size of *Eusirus* spp. could be estimated at 40-47 mm, while the size of further unidentified amphipod species was estimated at ranging between 11 and 19 mm. One stomach was completely empty. Almost all parasites found in the stomachs were trematodes, which occurred in 81% of the investigated haddock (**Table 34**). Two individual nematode parasites were found in a single stomach. One non-organic item was found in a haddock stomach, which was identified as a stone.

Beaked redfish: The stomachs of the beaked redfish contained little recognizable food items and one stomach was completely empty. In one stomach, recognizable food items included amphipods, copepods and other unidentifiable crustacean remains (**Figure 61***,* **Table 33**). The stomach further contained few small body parts indicating the presence of chaetognaths and fish. Trematode parasites were found in two of the three stomachs analysed.

Black seasnail: The stomach of the black seasnail contained the mesopelagic amphipod *Cyclocaris guilelmi,* an unidentifiable amphipod and other crustacean remains.

Fish food-web structure based on $\delta^{13}C$ and $\delta^{15}N$ stable isotope data

Muscle samples from 9 polar cod (4 from the SAS2021 Oden expedition, 5 from MOSAiC), 14 Atlantic cod, 38 haddock, one ice cod and four beaked redfish were analysed for $\delta^{13}C$, $\delta^{15}N$ and C:N ratio. The stable isotope composition of muscle from the remaining nine polar cod sampled during MOSAiC were sampled in another project (F. Schaafsma, unpublished data). Regarding $\delta^{13}C$ and $\delta^{15}N$ values, polar cod were separated from Atlantic cod, haddock and ice cod (**Figure 62**). The $\delta^{15}N$ values of polar cod were (with one exception) generally lower than in the other investigated fish species, with most values between 11 and 14 ‰ (mean 12.4 ‰; **Figure 63 B**, **Table 35**). The $\delta^{13}C$ values of all polar cod were lower than those of all other fish species analysed,
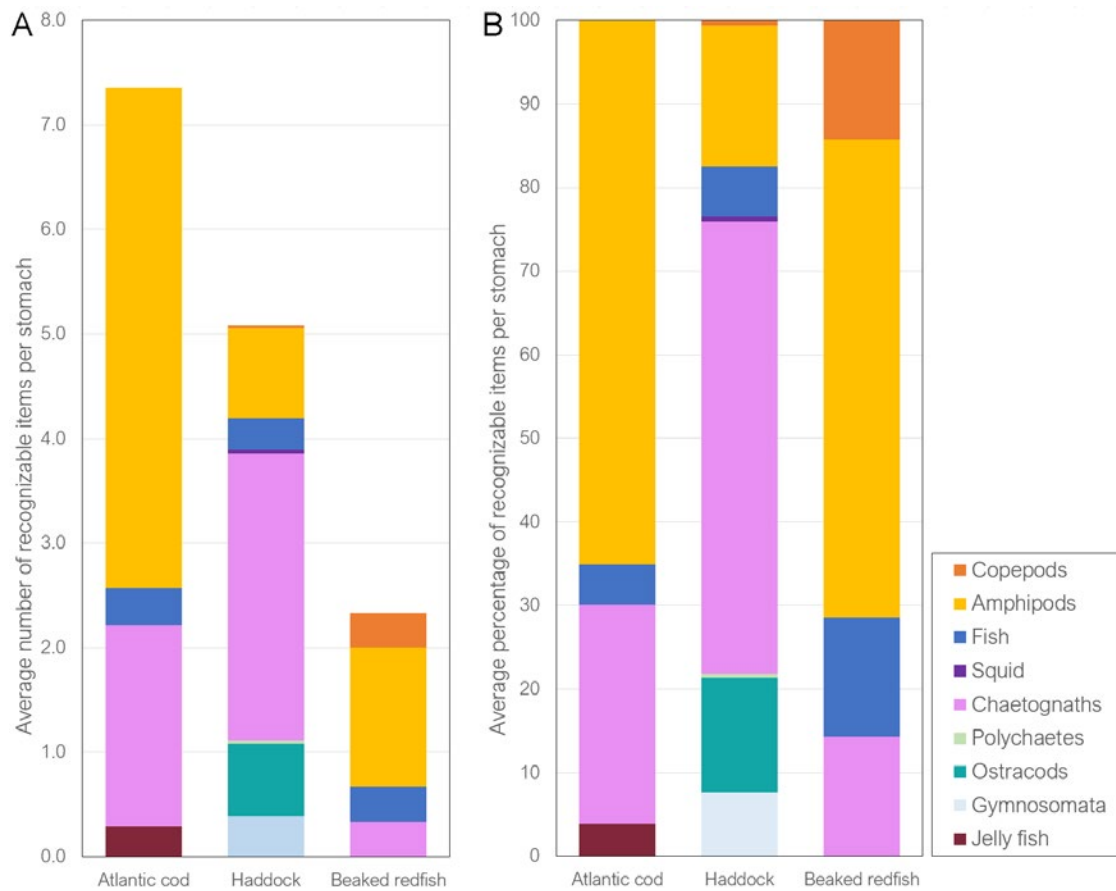
Figure 61: Average number of recognizable prey items found in the stomachs per species of Atlantic cod *Gadus morhua*, haddock *Melanogrammus aeglefinus*, and beaked redfish *Sebastes mentella*. (A) Prey items expressed in absolute numbers. (B) Prey items expressed as % of absolute number. Averages are based on the estimated Minimum Number of Individuals (MNI) of prey items per stomach.

Table 33: Summary of the recognizable prey items found in the stomach content of Atlantic cod *Gadus morhua*, haddock *Melanogrammus aeglefinus*, and beaked redfish *Sebastes mentella*. Results are expressed as the average Minimum Number of Individuals (MNI), the average percentage of the total minimum number of food items and the frequency of occurrence (FO) per predator species.

	Atlantic cod			Haddock			Beaked redfish		
	Average MNI	Average %	FO	Average MNI	Average %	FO	Average MNI	Average %	FO
Copepods	0.0	0	0	0.0	1	3	0.3	14	33
Amphipods	4.8	65	57	0.9	17	39	1.3	57	33
Fish	0.4	5	21	0.3	6	31	0.3	14	33
Squid	0.0	0	0	0.0	1	3	0.0	0	0
Chaetognaths	1.9	26	14	2.8	54	56	0.3	14	33
Polychaetes	0.0	0	0	0.0	1	3	0.0	0	0
Ostracods	0.0	0	0	0.7	14	19	0.0	0	0
Gymnosomata	0.0	0	0	0.4	8	8	0.0	0	0
Jellyfish	0.3	4	21	0.0	0	0	0.0	0	0

Table 34: Detailed overview of the recognizable prey items found in the stomach content of Atlantic cod *Gadus morhua*, haddock *Melanogrammus aeglefinus*, and beaked redfish *Sebastes mentella*. Results are expressed as the average Minimum Number of Individuals (MNI) and the frequency of occurrence (FO) per predator species.

	Atlantic cod		Haddock		Beaked redfish	
	Average MNI	FO	Average MNI	FO	Average MNI	FO
Copepod unidentified	0.0	0.0	0.0	2.8	0.3	33.3
Amphipod unidentified	0.2	21.4	0.4	27.8	1.0	33.3
Cyclocaris guilelmi	2.8	7.1	0.0	0.0	0.0	0.0
Eusirus holmii	0.1	7.1	0.1	8.3	0.3	33.3
Eusirus spp.	0.0	0.0	0.1	2.8	0.0	0.0
Themisto abyssorum	0.2	14.3	0.2	13.9	0.0	0.0
Themisto libellula	0.1	14.3	0.0	2.8	0.0	0.0
Themisto spp.	1.4	28.6	0.0	2.8	0.0	0.0
Fish unidentified	0.4	21.4	0.3	30.6	0.3	33.3
Squid unidentified	0	0	0.0	2.8	0	0
Chaetognath unidentified	1.0	14.3	2.8	55.6	0.3	33.3
Sagitta maxima	0.9	7.1	0	0	0	0
Gymnosomata unidentified	0	0	0.4	8.3	0	0
Polychaete unidentified	0	0	0.0	2.8	0	0
Ostracod unidentified	0	0	0.7	19.4	0	0

Table 35: Stable isoope (δ^{13}C, δ^{15}N) and C:N ratio in muscle tissue of polar cod, ice cod, Atlantic cod, haddock and beaked redfish. The data are shown as mean ± 1 standard deviation. n = total number of fish analysed in duplicate (replicate samples for each fish individual).

Species	n	δ^{13}C (‰)		δ^{15}N (‰)		C:N ratio	
		Mean	Range (min to max)	Mean	Range (min to max)	Mean	Range (min to max)
Polar cod	9	-24.3 ± 0.6	-25.4 to -23.8	12.4 ± 1.3	11.3 to 15.1	3.4 ± 0.1	3.3 to 3.5
Ice cod	1	-20.7 ± 0.3	-20.9 to -20.5	16.9 ± 0.2	16.8 to 17.0	3.1 ± 0.1	3.1 to 3.2
Atlantic cod	14	-21.2 ± 0.5	-22.3 to -20.4	13.0 ± 0.9	11.9 to 15.1	3.1 ± 0.1	3.1 to 3.3
Haddock	38	-20.3 ± 0.8	-23.1 to -18.8	13.3 ± 0.8	11.9 to 14.9	3.3 ± 0.4	3.0 to 5.7
Beaked redfish	4	-20.5 ± 0.1	-20.7 to -20.3	13.3 ± 0.9	12.2 to 14.4	3.2 ± 0.1	3.2 to 3.3

ranging between -25.4 and -23.8 ‰ (mean -24.3 ‰; **Figure 63 A**, **Table 35**). The δ^{13}C value range of polar cod was within the range of polar cod and zooplankton prey species from the CAO[134,135] and the ranges of zooplankton measured during the MOSAiC (-29 to -21 ‰; Katrin Schmidt, personal communication) and SAS-Oden expeditions (-29 to -21 ‰). The δ^{15}N values of polar cod were on average slightly higher than in

---

[134] Kohlbach D, et al. (2016) The importance of ice algae-produced carbon in the central Arctic Ocean ecosystem: Food web relationships revealed by lipid and stable isotope analyses. Limnology & Oceanography 61:2027–2044 [http://doi.org/10.1002/lno.10351]

[135] Kohlbach D, et al. (2017) Strong linkage of polar cod (*Boreogadus saida*) to sea ice algae-produced carbon: Evidence from stomach content, fatty acid and stable isotope analyses. Progress in Oceanography 152:62–74 [http://doi.org/10.1016/j.pocean.2017.02.003]

earlier measurements[136], and $\delta^{15}N$ values of Atlantic cod, haddock and beaked redfish were within the range reported from adjacent waters[137].

Atlantic cod, haddock and beaked redfish all clustered together in a group with most $\delta^{13}C$ values between -22 and -18 ‰, and $\delta^{15}N$ values between 12 and 15 ‰. One Atlantic cod caught during MOSAiC Leg 1 and one haddock caught during MOSAiC Leg 3 had intermediate $\delta^{13}C$ values between polar cod and the other fish species (**Figure 62**). The ice cod was similar to Atlantic cod, haddock and beaked redfish in $\delta^{13}C$ values, but had the highest $\delta^{15}N$ value of all fish investigated (16.9 ‰; **Table 35**).

There is often a positive correlation between $\delta^{15}N$ and $\delta^{13}C$ values in codfishes, which could in parts explain the significantly higher $\delta^{13}C$ values in the Atlantic fish compared to polar cod and prey organisms of the CAO. However, whether differences in $\delta^{15}N$ values between polar cod and the other species indicate a difference in the trophic position of the fish cannot be estimated at this point, because values from the trophic baseline (the $\delta^{15}N$ values of phytoplankton and ice algae) from the MOSAiC and SAS-Oden expedition and from potential areas of origin of the fish are not yet available. More importantly, the muscle tissue of the Atlantic fish may still have reflected the isotopic signal of the areas of origin of the fish. The mean half-life of $\delta^{13}C$ in muscle tissue of fish with 1-2 kg mass would be about 4 months at a temperature of 10 °C[138], and much longer at 1-3 °C. The $\delta^{13}C$ values of muscle tissue from most Atlantic cod and haddock match those from the Northern Norwegian Sea and Spitzbergen, where the source populations of these fish probably live. In these shelf waters, Atlantic cod and haddock rely predominantly on benthic prey, which has higher $\delta^{13}C$ values than the pelagic prey in the CAO and the Atlantic inflow region.

To replace this "benthic" signal, the fish would need to find sufficient food in the Arctic pelagic food web to not only cover their basic metabolism, but also to assimilate surplus carbon over many months. Hence, the near-absence of the $\delta^{13}C$ signature of the CAO food web in the muscle tissue samples of Atlantic cod, haddock and beaked redfish could be due to low tissue turnover rates in combination with zero or very low assimilation rates of carbon from the pelagic food web in the Arctic Ocean. The three Atlantic cod which were caught in the CAO had lower $\delta^{13}C$ in relation to $\delta^{15}N$ levels compared to the other Atlantic cod and haddock, which could indicate an influence of the pelagic food web of the CAO (**Figure 62**).

The C:N ratio is considered an indicator for lipid content, and thus for nutritious value and energetic condition of the fish, with higher values indicating higher nutritious value and better condition. The significantly higher C:N ratio of polar cod compared to the other investigated fish species reflects the known high lipid content of this key forage fish of the Arctic Ocean[139]. The mean C:N ratios of Atlantic cod and haddock sampled during MOSAiC, including fish from the CAO, were well within the reported range in adjacent waters[140], confirming the overall good condition of the fish sampled during the MOSAiC expedition (**Table 32**). Polar cod had generally higher C:N values (mean 3.4) than the other investigated fish species (mean 3.1 to 3.3; **Figure 63 C**, **Table 35**). Several

[136] Kohlbach D, et al. (2016) The importance of ice algae-produced carbon in the central Arctic Ocean ecosystem: Food web relationships revealed by lipid and stable isotope analyses. Limnology & Oceanography 61:2027-2044 [https://doi.org/10.1002/lno.10351]

[137] Ramsvatn S & Pedersen T (2012) Ontogenetic niche changes in haddock *Melanogrammus aeglefinus* reflected by stable isotope signatures, $\delta^{13}C$ and $\delta^{15}N$. Marine Ecology Progress Series 451:175-185 [https://doi.org/10.3354/meps09604]

[138] Thomas SM, Crowther TW (2015) Predicting rates of isotopic turnover across the animal kingdom: a synthesis of existing data. Journal of Animal Ecology 84:861-870 doi https://doi.org/10.1111/1365-2656.12326

[139] Marsh JM, Mueter FJ, Iken K, Danielson S (2017) Ontogenetic, spatial and temporal variation in trophic level and diet of Chukchi Sea fishes. Deep Sea Research Part II: Topical Studies in Oceanography 135: 78-94 doi https://doi.org/10.1016/j.dsr2.2016.07.010

[140] Linnebjerg JF, et al. (2016) Deciphering the structure of the West Greenland marine food web using stable isotopes ($\delta^{13}C$, $\delta^{15}N$). Marine Biology 163:230 [https://doi.org/10.1007/s00227-016-3001-0]

haddock individuals had very high C:N values of up to 5.7 (**Table 35**). Differences in $\delta^{13}C$ values, $\delta^{15}N$ values and C:N ratios were statistically significant between species (Kruskal-Wallis test; $\delta^{13}C$: $Chi^2=33.6$, $p<0.001$; $\delta^{15}N$: $Chi^2=11.3$, $p<0.05$; C:N: $Chi^2=25.0$, $p<0.001$). Within species, there were no statistically significant differences between expedition legs (Kruskal-Wallis test; $p>0.05$ for polar cod, Atlantic cod, haddock).
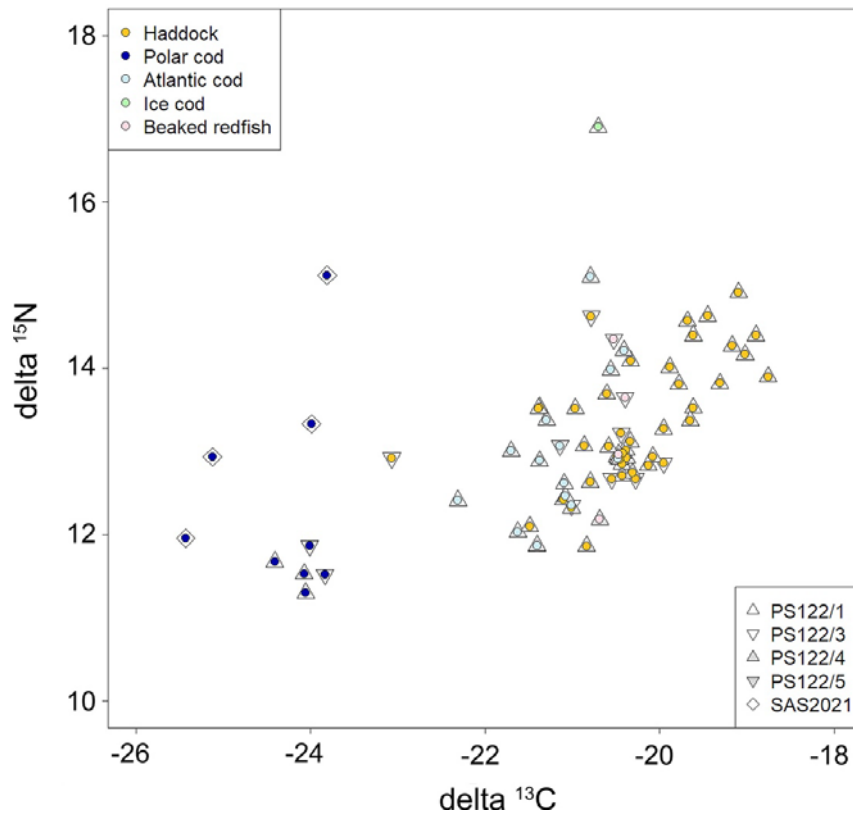
Figure 62: Scatter plot of $\delta^{15}N$ versus $\delta^{13}C$ from muscle tissue samples of haddock, polar cod, Atlantic cod, ice cod and beaked redfish sampled during the MOSAiC and SAS2021 Oden expeditions. Each value represents the mean of duplicate measurements from the same individual. PS122/1 – PS122/5 = MOSAiC expedition legs; SAS2021 = SAS-Oden expedition.
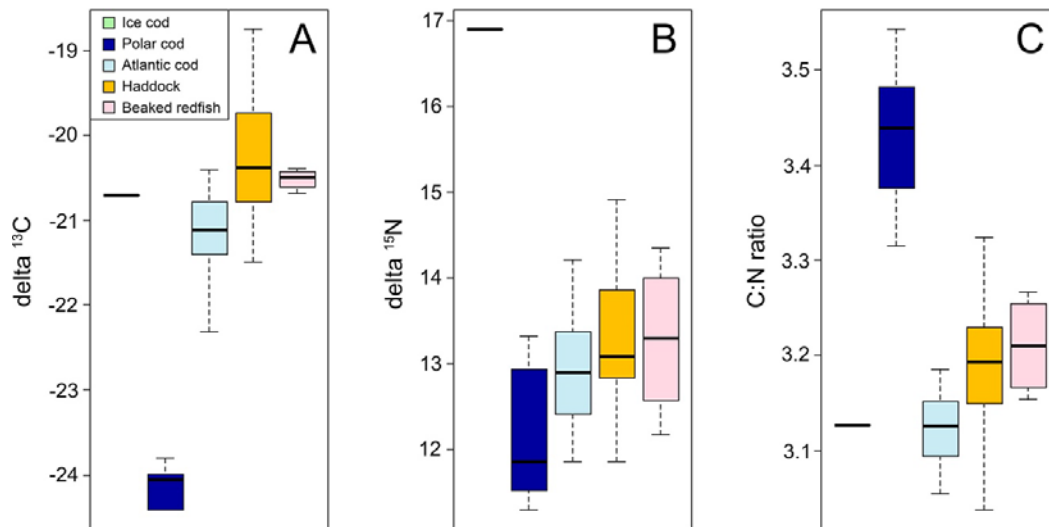
Figure 63: Comparison of value distributions in (A) δ¹³C, (B) δ¹⁵N, and (C) C:N ratios in muscle tissue of polar cod, ice cod, Atlantic cod, haddock and beaked redfish. Data are shown as standard box plots. The median value is indicated as a horizontal bar. The upper and lower margins of the "box" indicate the 25 and 75 percentiles, respectively, and the error bar shows 1.96 standard deviations. Outliers (extreme values) are not shown (see Table 35).

Fish food-web structure based on fatty acids composition

The total FA content in fish muscle was 2-3 times higher in *Paraliparis bathybius* and *Boreogadus saida* (ca. 7000-8000 ng FA mg^{-1} wet weight) than in the other species (ca. 2000-3000 ng FA mg^{-1} wet weight) (**Figure 64**, **Table 36**). The FA content in fish liver was ca. 50 times higher than that in muscle (**Figure 64**, **Table 37**). Altogether, 24 FAs were detected in fish muscle samples (**Table 36**) and 30 in fish liver samples (**Table 37**). All FAs detected in zooplankton samples except for 21:0 were also detected in the fish liver samples (see WP4).

Similar to the zooplankton, the dominant SAFAs in fish were C14:0 (myristic acid) and C16:0 (hexadecanoic acid), the dominant MUFAs were C16:1n-7 (palmitoleic acid) and C18:1n-9 (elaidic acid), and the dominant PUFAs were 20:5n-3 (eicosapentaenoic acid, EPA) and 22:6n-3 (docosahexaenoic acid, DHA), but in the fish livers also FA20:1n-11 (cis-11-eicosenoic acid) and 18:4n-3 (stearidonic acid) levels were high. Similar to zooplankton the content of bound FAs were higher in fish muscle than that of free FAs (**Figure 65**), and the same was found for fish liver (not shown). The content of PUFAs was generally higher than that of MUFAs, except for the only benthopelagic fish *Paraliparis bathybius*.

The PUFA/SAFA ratio is commonly used as an index for evaluating the nutritional value of a food item. In fish muscle the PUFA/SAFA for bound FAs was rather stable (ca. 0.8-1.0) for all species. For free FAs this ratio was highest in the ice-associated *Boreogadus saida* (2.5) (**Figure 66**). This also suggests that *Boreogadus saida* was in good condition and well-adapted to its extremely cold habitat (always below 0 °C). The other fish species were caught further down in the water column at temperatures of ca. 1 °C in the Atlantic water layer. The superiority of the nutritional value of *Boreogadus saida* is shown even more clearly when considering another indicator, the ω3/ω6 ratio. The ω3/ω6 ratio was about twice as high in *Boreogadus saida* than in all other species (**Figure 66**). The ω3/ω6 ratio of the three *Gadus morhua* individuals from the CAO (7.6 ± 2.1) did not differ significantly from the nine *Gadus morhua* individuals from the Atlantic inflow region (9.3

± 1.6), suggesting that the fish were in equal good condition. However, it should be kept in mind that the number of available fish individuals was very low.
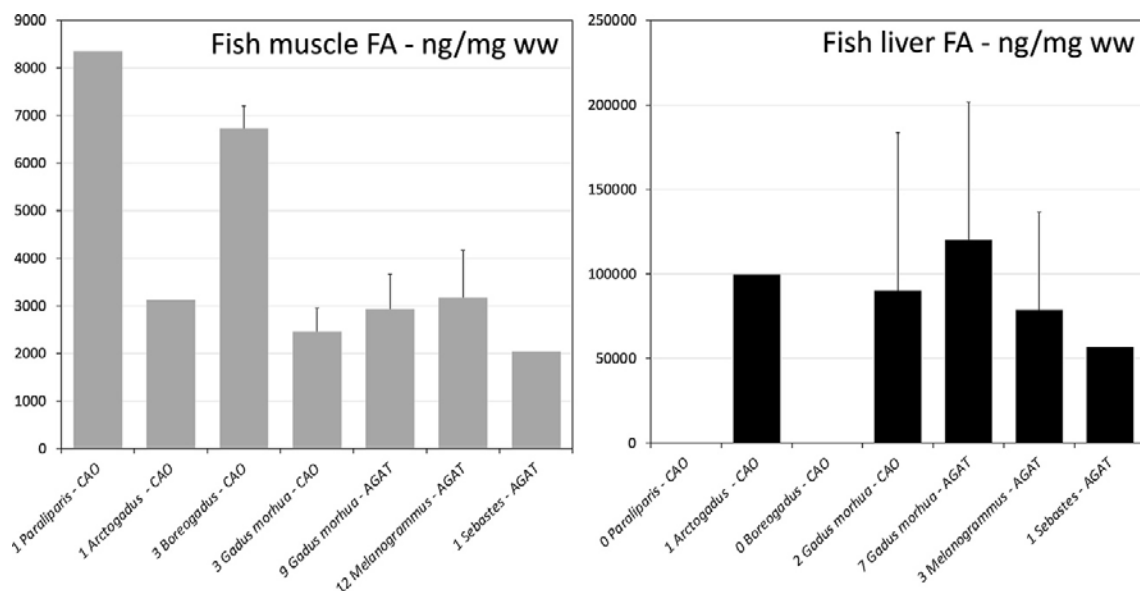
Figure 64: Mean content of fatty acids (FA) in fish muscle and fish liver samples. The number of replicate fish individuals of the same species is given before the fish name on the X-axis and the area where the fish was sampled is given after the fish name on the X-axis. CAO = Central Arctic Ocean, AGAT = Atlantic Gateway. The error bars represent 1 standard deviation of the mean.

Table 36: Results of the fatty acid (FA) analyses given as mean concentrations for the 24 fatty acids identified in fish muscle. b.d. = below detection limit. EPA = eicosapenta-enoic acid, DHA = docosahexaenoic acid. Data are shown as mean ± 1 standard deviation.

	Paraliparis bathybius CAO	Arctogadus glacialis CAO	Boreogadus saida CAO	Gadus morhua CAO	Gadus morhua AGAT	Melanogrammus aeglefinus AGAT	Sebastes mentella AGAT
Nr of individuals	n=1	n=1	n=3	n=3	n=9	n=12	n=1
Saturated fatty acids (SAFA) in ng mg^{-1} wet weight							
12:0	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.
13:0	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.
14:0	142	174	354 ± 92	139 ± 27	128 ± 22	154 ± 56	82
15:0	26	23	39 ± 10	20 ± 3	23 ± 3	44 ± 12	17
16:0	1528	939	2062 ± 121	719 ± 197	915 ± 201	982 ± 241	699
17:0	7	5	8 ± 1	5 ± 1	6 ± 2	11 ± 3	4
18:0	178	78	145 ± 21	72 ± 25	107 ± 32	119 ± 35	91
20:0	0.2	0.2	0.2 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	0.2 ± 0.1	0.1
21:0	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.
22:0	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.
24:0	0.2	0.2	0.3 ± 0.1	0.2 ± 0.0	0.2 ± 0.0	0.2 ± 0.1	0.1
Monounsaturated fatty acids (MUFA) in ng mg^{-1} wet weight							
14:1n-5	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.
16:1n-7	1041	284	617 ± 125	218 ± 28	125 ± 42	165 ± 115	125
17:1n-7	2.3	0.6	1.4 ± 0.3	0.6 ± 0.1	1.2 ± 0.5	3.2 ± 2.8	0.5
18:1n-9, trans	2000	200	373 ± 30	231 ± 38	207 ± 65	224 ± 125	131
18:1n-9, cis	704	116	235 ± 34	97 ± 23	88 ± 30	121 ± 72	55

20:1n-11	616	108	163 ± 33	87 ± 10	94 ± 40	93 ± 80	74
22:1n-9	68	18	41 ± 15	17 ± 2	18 ± 6	24 ± 12	18
24:1n-9	2.7	1.8	2.3 ± 0.8	1.5 ± 0.2	1.8 ± 0.6	2.1 ± 0.8	1.5
Polyunsaturated fatty acids (PUFA) in ng mg^{-1} wet weight							
16:4n-3	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.
18:3n-3	13	2	16 ± 9	5 ± 1	5 ± 2	5 ± 2	3
18:4n-3	59	19	74 ± 47	16 ± 13	29 ± 15	32 ± 27	8
20:3n-3	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.
20:5n-3 (EPA)	784	359	1191 ± 311	333 ± 60	353 ± 94	409 ± 85	206
22:6n-3 (DHA)	918	685	1258 ± 207	395 ± 131	711 ± 240	628 ± 238	443
18:2n-6	107	21	63 ± 9	27 ± 4	29 ± 5	28 ± 9	44
18:3n-6	1.0	0.8	2.4 ± 1.4	0.6 ± 0.2	0.5 ± 0.2	0.5 ± 0.2	0.5
20:2n-6	9	8	17 ± 4	8 ± 2	9 ± 3	13 ± 8	4
20:3n-6	1.0	0.5	1.1 ± 0.4	0.3 ± 0.2	0.5 ± 0.2	0.3 ± 0.1	0.3
20:4n-6	141	70	52 ± 14	57 ± 11	73 ± 21	104 ± 30	39
22:2n-6	8	11	11 ± 1	6 ± 1	7 ± 5	6 ± 4	3
Summary of concentrations in ng mg^{-1} wet weight							
Sum FA	8354	3123	6728 ± 477	2456 ± 498	2932 ± 734	3169 ± 2003	2049
Sum SAFA	1881	1219	2609 ± 124	956 ± 237	1180 ± 248	1311 ± 325	894
Sum MUFA	4433	729	1432 ± 166	653 ± 90	535 ± 175	632 ± 396	404
Sum PUFA	2040	1176	2687 ± 415	847 ± 176	1217 ± 344	1226 ± 328	751
Sum ω3	1774	1065	2540 ± 392	748 ± 179	1098 ± 324	1074 ± 302	660
Sum ω6	266	111	147 ± 23	99 ± 7	119 ± 2.6	151 ± 42	91

Table 37: Results of the fatty acid analyses given as mean concentrations for the 30 fatty acids identified in fish liver. Liver samples of *Boreogadus saida* and *Paraliparis bathybius* were not available. b.d. = below detection limit. EPA = eicosapentaenoic acid, DHA = docosahexaenoic acid. Data are shown as mean ± 1 standard deviation.

	Arctogadus glacialis CAO	*Gadus morhua* CAO	*Gadus morhua* AGAT	*Melanogrammus aeglefinus* AGAT	*Sebastes mentella* AGAT
Nr of individuals	n=1	n=2	n=7	n=3	n=1
Saturated fatty acids (SAFA) in ng mg^{-1} wet weight					
12:0	833	170 ± 62	307 ± 77	111 ± 15	98
13:0	114	66 ± 33	92 ± 63	63 ± 37	12
14:0	5664	4070 ± 3290	5947 ± 4301	3417 ± 2326	1216
15:0	398	242 ± 159	426 ± 329	349 ± 207	173
16:0	13249	10424 ± 8001	13685 ± 9837	10180 ± 6213	6586
17:0	131	144 ± 134	239 ± 191	258 ± 179	145
18:0	1252	1915 ± 2064	2653 ± 2742	2335 ± 1419	2078
20:0	35	12 ± 9	25 ± 19	25 ± 12	21
21:0	b.d.	b.d.	b.d.	b.d.	b.d.
22:0	0.5	1.3 ± 1.4	1.3 ± 1.2	1.3 ± 1.1	0.4
24:0	3.2	1.5 ± 1.1	4.8 ± 3.2	2.8 ± 1.2	1.9
Monounsaturated fatty acids (MUFA) in ng mg^{-1} wet weight					
14:1n-5	351	314 ± 304	289 ± 255	229 ± 238	45
16:1n-7	17251	13307 ± 11470	12188 ± 10274	9232 ± 8580	3567
17:1n-7	248	178 ± 154	293 ± 259	378 ± 296	179

18:1n-9, trans	14386	14461 ± 14407	12777 ± 8943	11827 ± 9226	7242
18:1n-9, cis	2895	5924 ± 6681	5317 ± 5509	5192 ± 4332	2733
20:1n-11	16996	12867 ± 14172	15084 ± 11202	7813 ± 7205	5651
22:1n-9	12423	4341 ± 4888	9010 ± 6974	2889 ± 2654	4097
24:1n-9	857	314 ± 282	662 ± 415	376 ± 72	296
Polyunsaturated fatty acids (PUFA) in ng mg^{-1} wet weight					
16:4n-3	41	69 ± 65	37 ± 29	32 ± 33	8
18:3n-3	331	574 ± 692	825 ± 863	468 ± 492	349
18:4n-3	2003	4344 ± 5525	3416 ± 4004	2905 ± 3002	1006
20:3n-3	114	178 ± 180	171 ± 172	111 ± 77	118
20:5n-3 (EPA)	3353	5575 ± 7465	7956 ± 7116	6943 ± 4563	7308
22:6n-3 (DHA)	1892	5422 ± 7347	10386 ± 6039	8217 ± 4246	5523
18:2n-6	916	1230 ± 1388	1813 ± 1862	1238 ± 928	1016
18:3n-6	221	109 ± 128	130 ± 156	70 ± 80	37
20:2n-6	845	1721 ± 2312	4855 ± 4393	1775 ± 1650	2660
20:3n-6	99	78 ± 74	107 ± 129	52 ± 47	65
20:4n-6	869	922 ± 728	1813 ± 1401	1203 ± 410	1339
22:2n-6	2097	1203 ± 1540	9666 ± 7868	1170 ± 380	3186
Summary in ng mg^{-1} wet weight					
Sum FA	99867	90176 ± 93433	120174 ± 81602	78867 ± 57818	56758
sum SAFA	21679	17046 ± 13631	23380 ± 17212	16742 ± 10299	10332
Sum MUFA	65408	51706 ± 52358	55619 ± 41096	37938 ± 32351	23810
Sum PUFA	12780	21424 ± 27443	41175 ± 25016	24186 ± 15547	22616
Sum ω3	7734	16162 ± 21274	22791 ± 17740	18677 ± 12254	14312
Sum ω6	5046	5263 ± 6170	18383 ± 12418	5509 ± 3305	8304
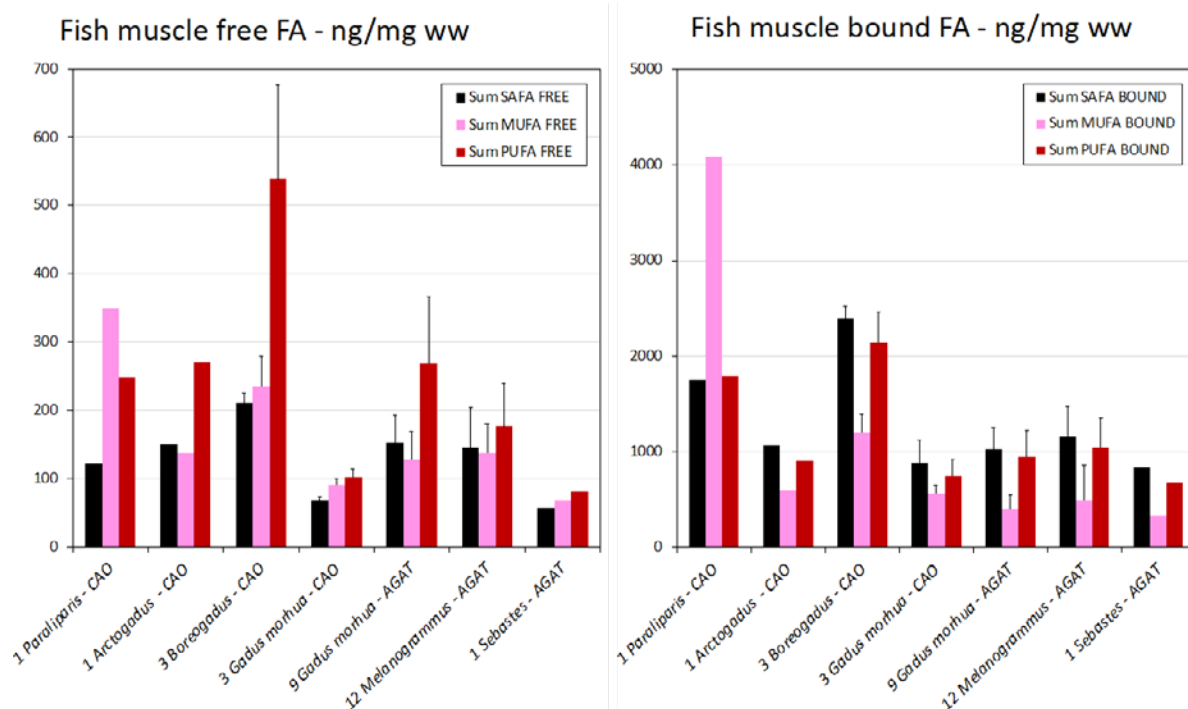
Figure 65: Mean contents of saturated fatty acid (SAFA), monounsaturated fatty acid (MUFA), and polyunsaturated fatty acid (PUFA) in the fish muscle samples analysed, to

the left free FA and to the right bound FA. The number of replicate fish individuals of the same species is given before the fish name on the X-axis and the area where the fish was sampled is given after the fish name on the X-axis. CAO = Central Arctic Ocean, AGAT = Atlantic Gateway. The error bars represent 1 standard deviation of the mean.
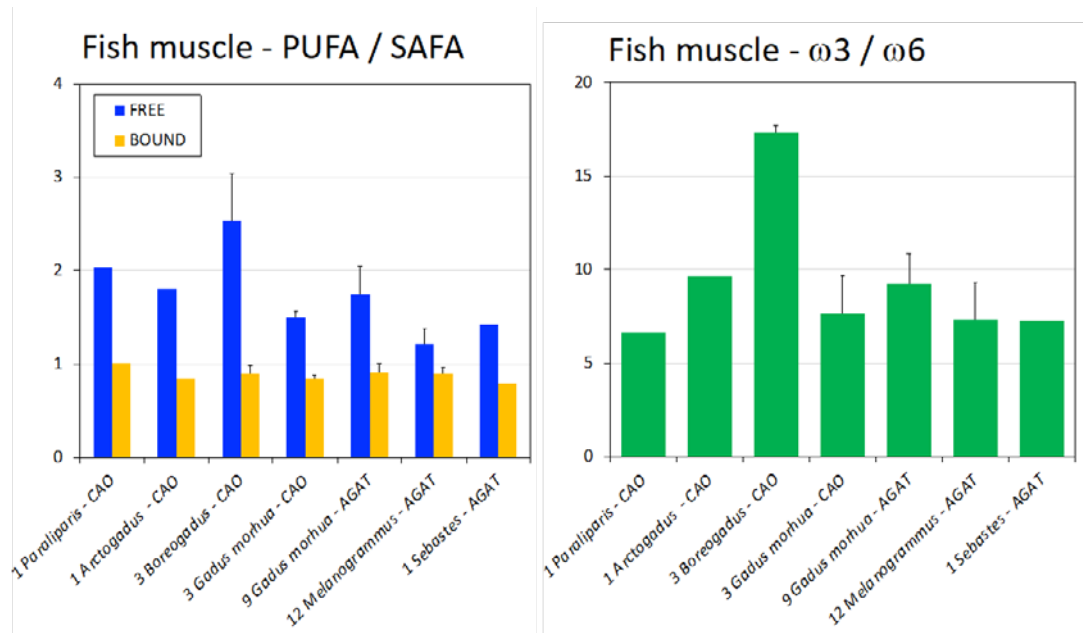
Figure 66: Indicators of nutritional value in fish muscle. To the left the ratio of polyunsaturated fatty acids (PUFA) to saturated fatty acids (SAFA), showing that this ratio was rather stable for bound FAs for all species, but variable for free FAs among species. To the right the $\omega 3/\omega 6$ ratio of the PUFAs based on total FAs (the sum of free and bound FAs), showing that this ratio was twice as high for polar cod _Boreogadus saida_ than for the other species. The number of replicate fish individuals of the same species is given before the fish name on the X-axis and the area where the fish was sampled is given after the fish name on the X-axis. CAO = Central Arctic Ocean, AGAT = Atlantic Gateway. The error bars represent 1 standard deviation of the mean.
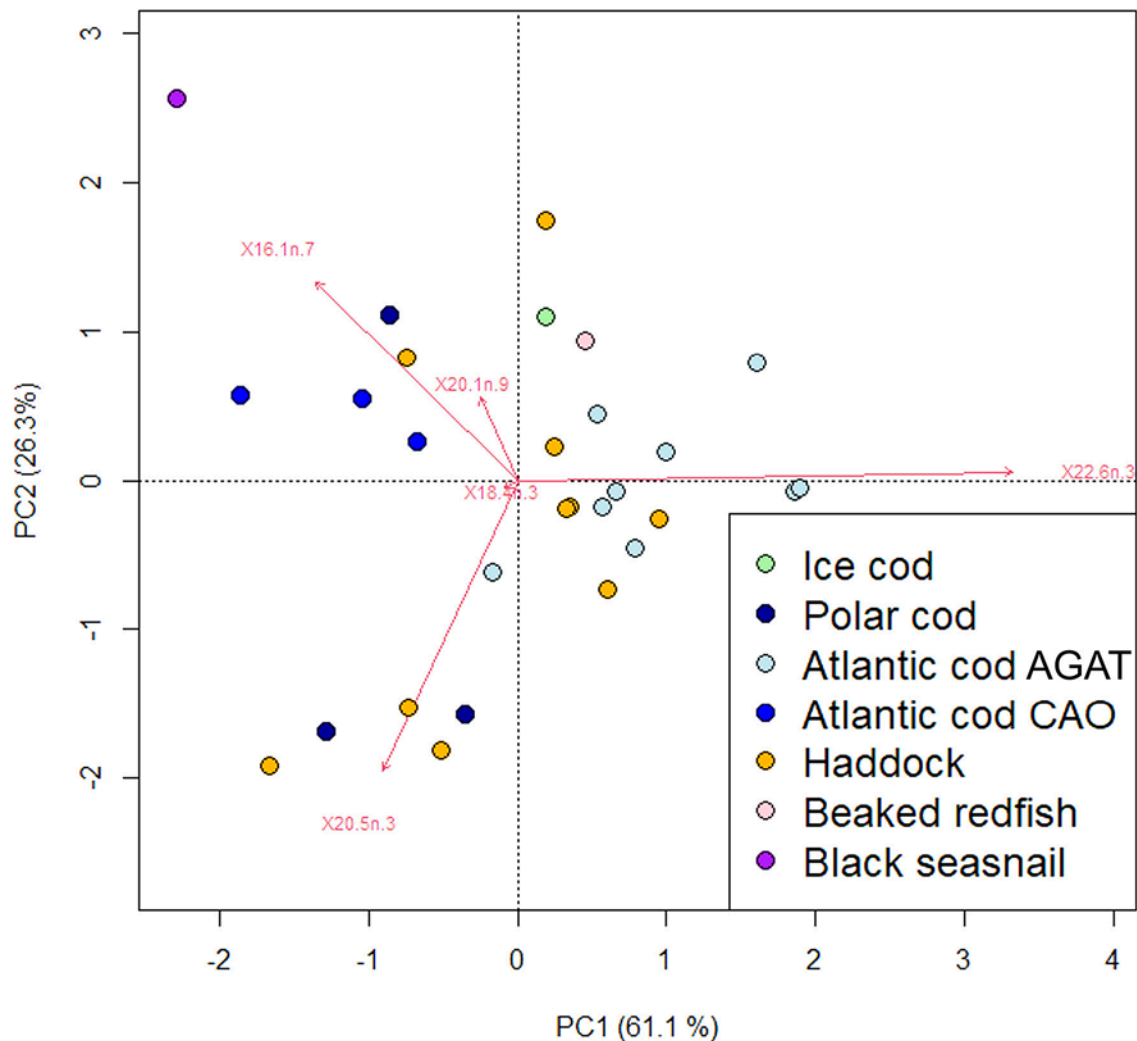
Figure 67: PCA biplot of the relative contribution of the five fatty acid trophic markers 16:1n-7, 20:5n-3, 18:4n-3, 22:6n-3 and 20:1n-9 in ice cod, polar cod, Atlantic cod, haddock, beaked redfish and black seasnail. CAO = Central Arctic Ocean, AGAT = Atlantic Gateway.

Fish age based on otolith increments

Within the WP07 project otoliths from 65 fish sampled during MOSAiC Legs 2-4 and the SAS-Oden expedition were analysed. These 65 otolith analyses exclude the six Atlantic cod (3 from Leg 1 and 3 from Leg 3) and the single ice cod (Leg 1) for which the data have previously been published[141]. However, the latter seven fish are also part of the SC07 project, so that the final data set consists of ages of 72 pelagic and sympagic fish of four species.

The 14 Atlantic cod were between 2 and 14 years old, most of them were 5-6 years old (mean 6.6 years; **Table 38**). The 38 haddock were between 3 and 15 years old (mean 5.8 years; **Table 38**). The majority of the haddock (22 fish) were 3 or 4 years old (**Figure 68**). The 15 polar cod sampled during MOSAiC and with otoliths allowing age reading were between 2 and 3 years old (mean 2.1 years; unpublished data H. Flores).

---

[141] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

The four polar cod sampled during the SAS-Oden expedition were between 1 and 3 years old (mean 2.0 years; **Table 38**).

Table 38: Summary of the results of fish age readings from otoliths of fish sampled during the MOSAiC and SAS-Oden expeditions. Data are shown as mean ± 1 standard deviation.

Fish species	Number of fish caught	Number of fish aged	Total length (cm)	Age (years)	Age range (years)
Atlantic cod	14	14	53.6 ± 12.8	5.8 ± 2.7	2-14
Haddock	38	38	51.0 ± 17.0	6.6 ± 4.3	3-15
Ice cod	1	1	33.2	4.0 ± 0.0	-
Polar cod (SAS-Oden)	4	4	13.2 ± 2.4	2.0 ± 0.8	1-3
Subtotal	75	72			
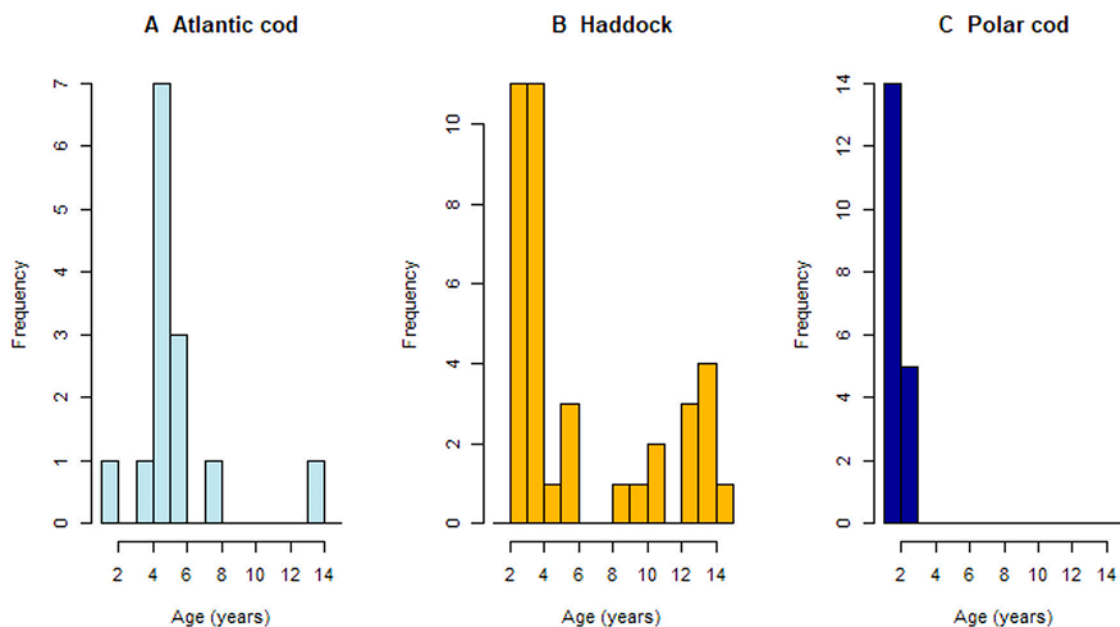
Figure 68: Frequency distributions of fish age at time of death determined from otolith increment readings. (A) Atlantic cod. (B) Haddock. (C) Polar cod. The graph for polar cod includes unpublished data from the MOSAiC expedition (data owner Hauke Flores, AWI).

Otolith δ^{13}C stable isotope analysis

Altogether, 574 samples from 67 fish (14 Atlantic cod, 34 haddock, 1 ice cod and 18 polar cod) were analysed for δ^{13}C stable isotope composition from otolith increment samples to assess the potential of this measurements to estimate field metabolic rate (FMR).

FMR is inversely related to δ^{13}C values in otolith samples. The data show a negative correlation of δ^{13}C values with ambient temperature in Atlantic cod and haddock, but not in polar cod (**Figure 69 A**). This could indicate that polar cod can maintain a constant field metabolic rate even at very low temperatures. The range in δ^{13}C values of ice cod falls within the range of haddock, but there are too few data points to infer a trend with temperature. The δ^{13}C values cannot be used to infer on potential differences in FMR

between polar cod versus Atlantic cod, haddock and ice cod, because they had dwelled in different water masses (Polar Surface Water and Atlantic Water, respectively). Comparing the two Atlantic species, otolith samples from haddock had generally higher $\delta^{13}C$ values than most Atlantic cod samples, potentially indicating overall lower metabolic rates in haddock than in Atlantic cod. In the life-history development of $\delta^{13}C$ values, this difference is even more evident (**Figure 69 B**). Most Atlantic cod have lower $\delta^{13}C$ values than haddock throughout their life history, which is potentially related to higher FMRs. This could mean that Atlantic cod can maintain a higher metabolic activity compared to haddock even at low temperatures. This observation may be important to predict differences in the physiological limits of the northern range expansion of the two species.

The estimation of FMR requires further complex processing steps, including the correction for individual biomass and shifts in the environmental baseline $\delta^{13}C$ values due to long-

term trends or differences between water masses. However, the preliminary raw data indicate that FMR can be estimated from $\delta^{13}C$ values at least in haddock and Atlantic cod, and potentially in polar cod. Such estimates can be valuable to assess the potential of these fish species to survive in the CAO for longer time.
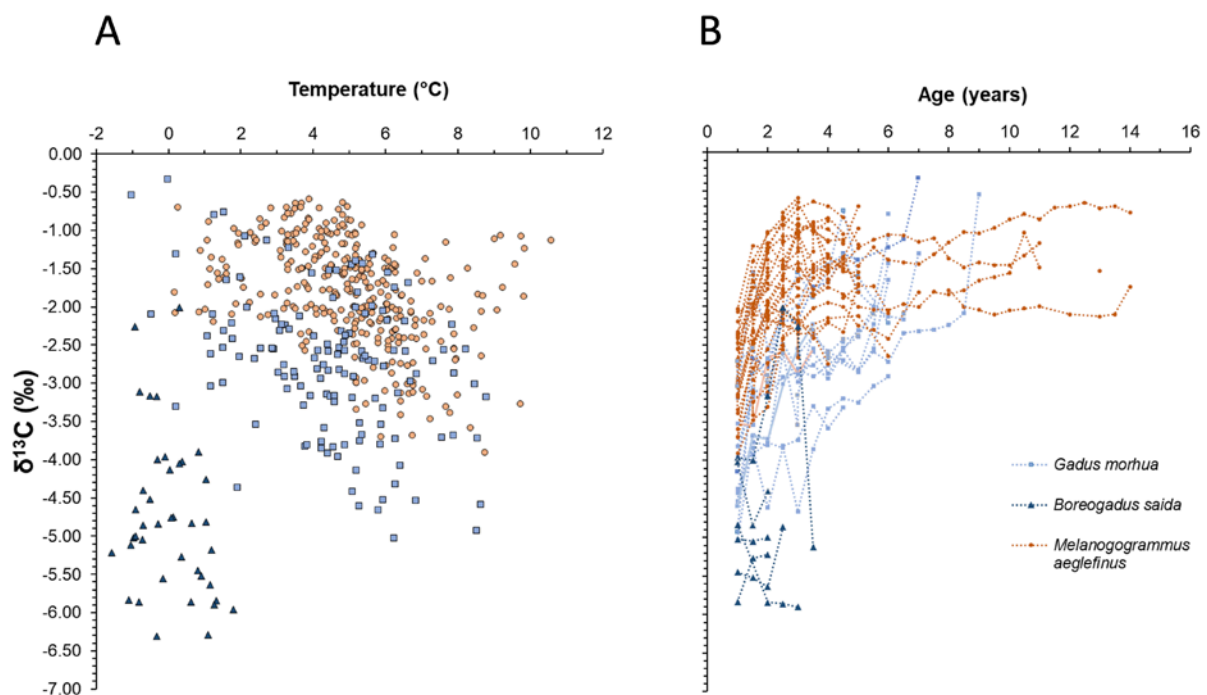
Figure 69: Results of stable isotope analysis ($\delta^{13}C$) from otolith increment samples of Atlantic cod *Gadus morhua*, polar cod *Boreogadus saida*, and haddock *Melanogrammus aeglefinus*. (A) Reconstructed ambient temperature. (B) Individual life histories.

Temperature reconstruction from otolith $\delta^{18}O$ stable isotope data

After quality check of the isotopic measurements, the ambient temperature at age in 540 of the 574 otolith increment samples analysed from Atlantic cod, haddock, ice cod and polar cod was estimated. The temperature reconstructions were based on assumed $\delta^{18}O$ value ranges of Atlantic water in Atlantic cod and haddock. Temperature reconstructions were not performed for ice cod because its region of origin, and therefore the $\delta^{18}O$ value ranges of ambient water in its early years of life, were not known. As polar cod migrate between Atlantic Water and Polar Surface Water, the $\delta^{18}O$ value of their ambient water was modelled as a mix between these two water masses. The mean reconstructed ambient temperatures from all increments were about 4 °C in Atlantic cod, 5 °C in

haddock, and close to 0 °C in polar cod (**Table 39**). The three Atlantic cod sampled in the North Pole region already published[142] indicated different migration dispersal patterns, including origins from both warmer north Atlantic waters (fish individual FR_10027) and Arctic waters (fish individuals FR_10002 and FR_10003) (**Figure 70 A**). The other 11 Atlantic cod individuals sampled in the Atlantic inflow region at the Yermak Plateau originated from warmer waters between 6 and 9 °C and indicated exposure to Arctic temperatures between -1 and 3 °C in their last year (**Figure 70 A**). Life-history temperature reconstructions of haddock caught at the Yermak Plateau and in the Fram Strait showed a similar pattern as the temperature reconstructions of Atlantic cod from the same region (**Figure 70 A, B**).

In eight polar cod sampled during MOSAiC, measurements of three or more increments were available for life-history temperature reconstructions. There was no common pattern in reconstructed ambient temperature during their lifespans (**Figure 70 C**). Five of the polar cod appeared to have dwelled in waters between about -1 and 0.5 °C, indicating that they remained in high-Arctic waters during their entire life. In three polar cod caught in the Yermak Plateau region, increments showed temperatures >1 °C at least during parts of their lifetime, indicating warmer waters, e.g., in the Barents Sea (**Figure 70 C**).

Table 39: Summary statistics of ambient temperature reconstructions (°C) in otolith increments from Atlantic cod, haddock, polar cod and ice cod. The data are shown as mean ± standard deviation. n = number of increment samples.

Species	n	Core		Outer increment		All increments	
		Mean	Range	Mean	Range	Mean	Range
Atlantic cod	162	5.7 ± 2.5	1.2 to 8.8	1.6 ± 1.9	-1.0 to 5.4	4.3 ± 2.1	-1.8 to 8.8
Haddock	329	6.9 ± 1.2	5.1 to 9.7	1.9 ± 1.3	0.2 to 6.4	5.1 ± 1.9	0.2 to 10.6
Polar cod	40	-0.1 ± 0.7	-1.0 to 0.8	-0.1 ± 1.0	-1.6 to 1.8	0.0 ± 0.9	-1.6 to 8
Subtotal	540						

Stock structure based on population genetics

Atlantic cod (*Gadus morhua*)

Atlantic cod is a commercially exploited demersal marine fish species. It was assumed to be distributed along the coasts and continental shelves of the North Atlantic Ocean as far north as the Svalbard area[143]. Atlantic cod is managed in its northern range in ICES units 1, 2a and 2b. This boreal generalist species is generally highly mobile[144], and was within the present FWC (SC03) discovered in the CAO, increasing its distribution range with about four latitudinal degrees northward (from 82 to 86 °N)[145]. The largest known stock of Atlantic stock is the migratory Northeast Arctic Cod (NEAC), which is distributed along the Norwegian coast, the Barents Sea and off Svalbard. The more stationary coastal cod (CC) spawn together with NEAC in the Lofoten area and several other areas along the

[142] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

[143] Haug T, et al. (2017) Future harvest of living resources in the Arctic Ocean north of the Nordic and Barents Seas: A review of possibilities and constraints. Fisheries Research 188:38-57 [https://doi.org/10.1016/j.fishres.2016.12.002]

[144] Frainer A, et al. (2017) Climate-driven changes in functional biogeography of Arctic marine fish communities. PNAS 114:12202–12207 [https://doi.org/10.1073/pnas.1706080114]

[145] Snoeijs-Leijonmalm P, et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

Norwegian coast. Allele frequencies at the pantophysin (*Pan I*) gene locus differ between ecotypes, with high frequencies of the *Pan I AA* genotype characteristic of CC and the *Pan I BB* genotype predominating in NEAC[146].

[146] Spotowitz L, et al. (2022) New evidence for the establishment of coastal cod *Gadus morhua* in Svalbard fjords. Marine Ecology Progress Series 696:119-133 [https://doi.org/10.3354/meps14126]
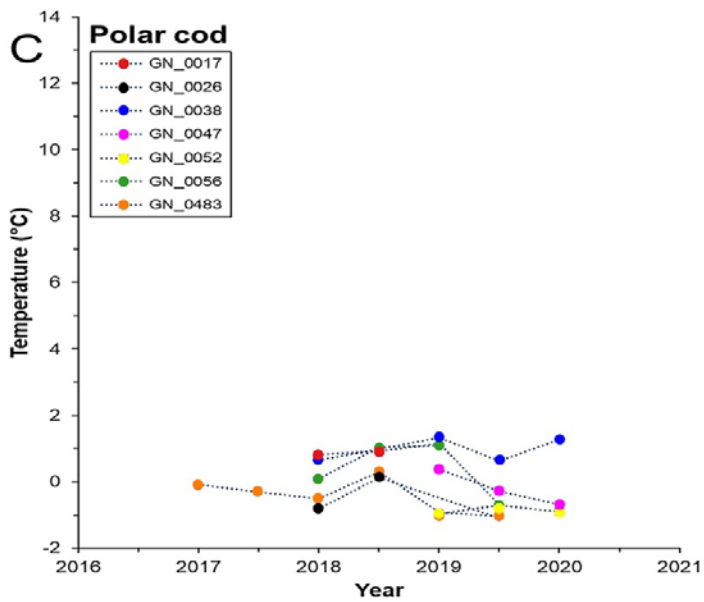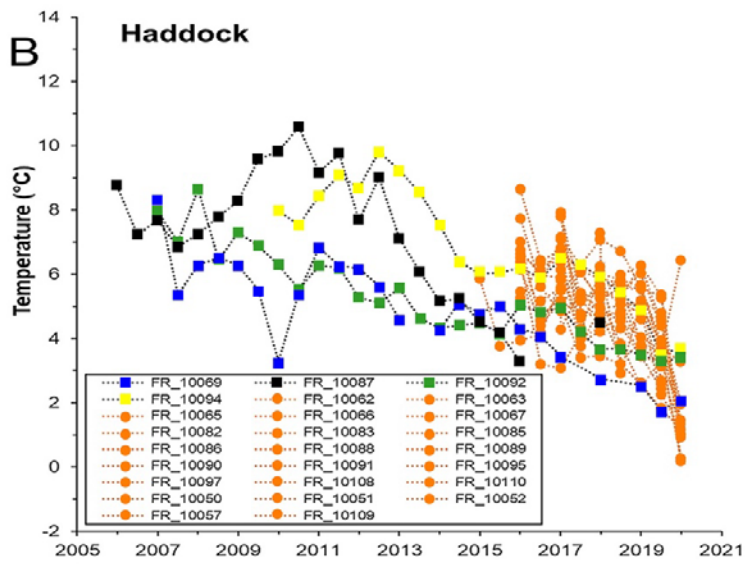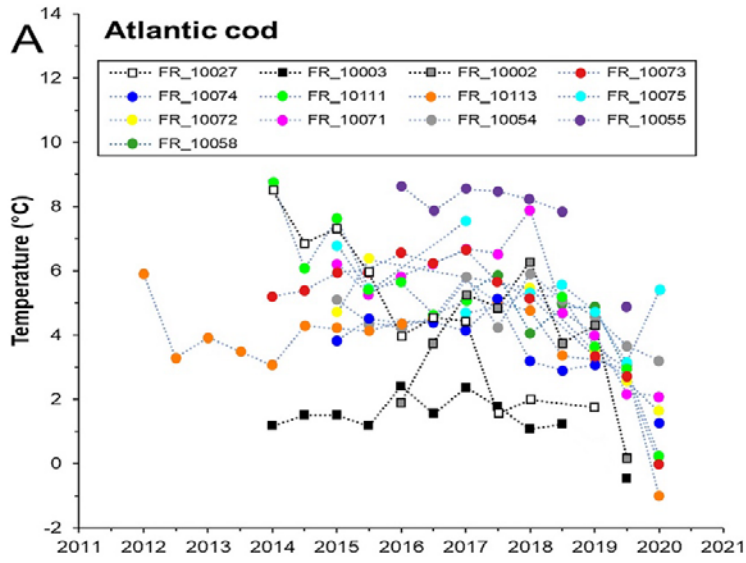
Figure 70: Life-history temperature reconstruc-tions based on δ¹⁸O data from otolith increment samples of (A) Atlantic cod, (B) haddock, and (C) polar cod. Temperature curves highlighted in black refer to fish caught in the North Pole region in A and to haddock ≥ 6 years in B. The legends ('FR_10xxx', 'GN_00xx') indicate sample labels of individual fish. In (B) all fish individuals younger than 6 years have the same legend (orange circles) to show the general downward trend in temperature between 2015 and 2021 for these fish as a group; single fish cannot be clearly distinguished in the graph due to the high number of fish in this age group.

All 14 Atlantic cod specimens sampled by the EFICA Consortium in 2019-2020 were successfully genotyped at the *Pan I* locus. Individuals showed either the homozygous *BB* genotype (n=13; Atlantic Gateway and CAO) or the heterozygous *AB* genotype (n=1; Atlantic Gateway) (**Figure 71**, **Table 40**), including re-analyses of the six Atlantic cod analysed earlier with the same results[147]. These genotypes most likely match with the North East Atlantic cod (NEAC) stock[148]. The recent colonisation of the CAO by the migratory NEAC ecotype fits with a gradual shift of the boreal fauna into northern waters[149]. Of great interest is that breeding populations of NCC and NEAC have colonised waters off Svalbard[150]. Classical knowledge states that the northern breeding grounds of NEAC and NCC are off Lofoten (Norway), with feeding grounds of NEAC occupying the Barents Sea.

Extension of the natural range of Atlantic fish species poleward as recorded in the Barents Sea fits with the gradual change in physical oceanography due to global warming[151]. At first sight, the records of Atlantic cod in the central Amundsen Basin provide no evidence of a recent expansion of this species into the CAO with climate change. The circulation of the comparatively warm and salty Atlantic water in the CAO [152]has existed for at least 14 million years[153], and the absence of Atlantic cod from the CAO in the past cannot be proven. However, it cannot be excluded that connectivity between the CAO and Atlantic species (Atlantic cod, haddock, beaked redfish, and probably others) has only recently emerged as a result of "Atlantification" of the Arctic shelf seas.

Haddock (*Melanogrammus aeglefinus*)

Haddock is a commercially exploited demersal marine fish species distributed along the coasts and continental shelves of the North Atlantic Ocean. Haddock is managed in its northern range in ICES units 1, 2a and 2b. The Similar to Atlantic cod, haddock is a highly mobile species[154]. Historically, its natural range extends all over the Barents Sea between Northern Norway, the White Sea, Novaya Zemlya, and the west and southeast

[147] Snoeijs-Leijonmalm P. et al. (2022) Unexpected fish and squid in the central Arctic deep scattering layer. Science Advances 8:eabj7536 [https://www.science.org/doi/10.1126/sciadv.abj7536]

[148] Johansen T, et al. (2018) "Real-time" genetic monitoring of a commercial fishery on the doorstep of an MPA reveals unique insights into the interaction between coastal and migratory forms of the Atlantic cod. ICES Journal of Marine Science 75:1093–1104 https://doi.org/10.1093/icesjms/fsx224

[149] Haug T, et al. (2017) Future harvest of living resources in the Arctic Ocean north of the Nordic and Barents Seas: A review of possibilities and constraints. Fisheries Research 188:38–57 https://doi.org/10.1016/j.fishres.2016.12.002

[150] Spotowitz L, et al. (2022) New evidence for the establishment of coastal cod *Gadus morhua* in Svalbard fjords. Marine Ecology Progress Series 696:119–133 (https://doi.org/10.3354/meps14126]

[151] Fossheim M et al. (2015) Recent warming leads to a rapid borealization of fish communities in the Arctic. Nature Climate Change 5:673–678 [https://doi.org/10.1038/NCLIMATE2647]

[152] Rudels B. Arctic Ocean circulation, in Encyclopaedia of Ocean Sciences, J. H. Steele, Ed. (Academic Press, 2019), pp. 211–225

[153] Chen TY (2012) Variations of North Atlantic inflow to the Central Arctic Ocean over the last 14 million years inferred from hafnium and neodymium isotopes. Earth and Planetary Science Letters 353:82–92 [https://doi.org/10.1016/j.epsl.2012.08.012]

[154] Frainer A, et al. (2017) Climate-driven changes in functional biogeography of Arctic marine fish communities. PNAS 114:12202–12207 [https://doi.org/10.1073/pnas.1706080114]

coast of Svalbard[155,156]. The Barents Sea is a major feeding ground for juvenile and adult haddock during the summer months[157].
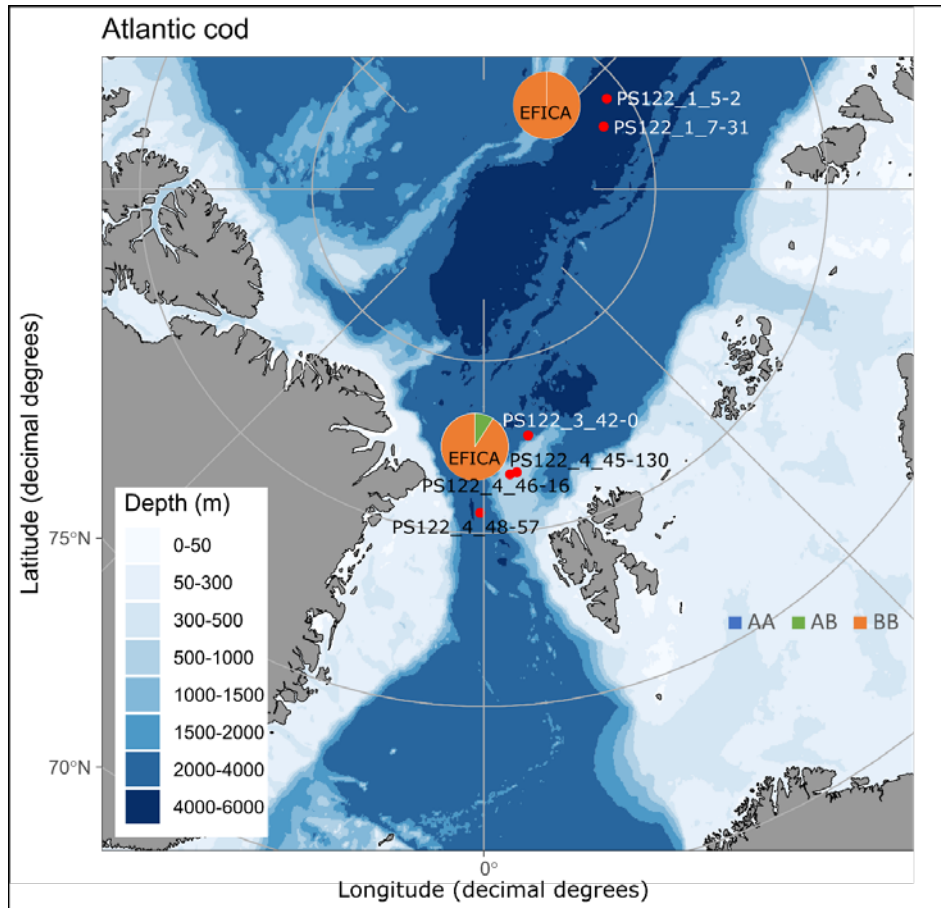
Figure 71: Map of the Arctic Ocean with Atlantic cod *Gadus morhua* sampling sites (red bullets) showing associated Pan I genotypes. The pie symbols show the distribution of the Pantophysin I genotype.

Table 40: Metadata and Pantophysin I genotypes (AB or BB) of the 14 Atlantic cod *Gadus morhua* specimens collected in the Central Arctic Ocean (CAO) and the Atlantic Gateway (AGAT) during the MOSAiC expedition.

Sample code	Tissue code	Station	Area	Latitude (°N)	Longitude(°E)	Marker	Genotype
FR_10003	GN_0070	1_5-2	CAO	85.54	126.44	Pan I	BB
FR_10002	GN_0071	1_5-2	CAO	85.54	126.44	Pan I	BB
FR_10027	GN_0087	1_7-31	CAO	86.05	117.68	Pan I	BB
FR_10054	GN_0158	3_42_0	AGAT	82.72	10.22	Pan I	BB
FR_10055	GN_0163	3_42_0	AGAT	82.72	10.22	Pan I	BB
FR_10058	GN_0178	3_42_0	AGAT	82.39	8.34	Pan I	BB

---

[155] Quintela M, et al. (2021) Distinct genetic clustering in the weakly differentiated polar cod, *Boreogadus saida* Lepechin, 1774 from East Siberian Sea to Svalbard. Polar Biology 44:1711–1724 [https://doi.org/10.1007/s00300-021-02911-7]

[156] https://www.fishbase.se

[157] Olsen E, et al. (2010) Cod, haddock, saithe, herring, and capelin in the Barents Sea and adjacent waters: A review of the biological value of the area. ICES Journal of Marine Science 67:87-101 [https://doi.org/10.1093/icesjms/fsp229]

FR_10071	GN_0006	4_45-130	AGAT	81.71	6.66	Pan I	BB
FR_10072	GN_0007	4_45-130	AGAT	81.71	6.66	Pan I	BB
FR_10073	GN_0009	4_45-130	AGAT	81.71	6.66	Pan I	AB
FR_10074	GN_0621	4_45-130	AGAT	81.71	6.66	Pan I	BB
FR_10075	GN_0011	4_46-16	AGAT	81.68	5.26	Pan I	BB
FR_10111	GN_0611	4_48-57	AGAT	80.59	-0.62	Pan I	BB
FR_10113	GN_0623	4_48-57	AGAT	80.59	-0.62	Pan I	BB
FR_10114	GN_0627	4_48-57	AGAT	80.59	-0.62	Pan I	BB

All 38 haddock specimens sampled by the EFICA Consortium in the Atlantic inflow region during the MOSAiC expedition (**Figure 72**) were successfully genotyped at 98 SNP markers (**Table 41**). Genetic diversity was relatively low with observed and expected diversity (H_o and H_e) values of 0.18 and 0.19, respectively, and an inbreeding coefficient (F_{IS}) of 0.02. From preliminary analyses it is suspected that all haddocks sampled belong to the European Arctic cluster ("Northeast Arctic" cluster)[158]. This result would be expected because the Northeast Arctic Ocean haddock stock is geographically closest to the Yermak Plateau area. presence of haddock in the Yermak Plateau region confirms the northward expansion, similarly to the northward range extension of other species such as Atlantic cod and mackerel[159].

---

[158] Berg PR, et al. (2021) Genetic structuring in Atlantic haddock contrasts with current management regimes. ICES Journal of Marine Science 78:1–13 [https://doi.org/10.1093/icesjms/fsaa204]

[159] Haug T, et al. (2017) Future harvest of living resources in the Arctic Ocean north of the Nordic and Barents Seas: A review of possibilities and constraints. Fisheries Research 188:38-57 [https://doi.org/10.1016/j.fishres.2016.12.002]
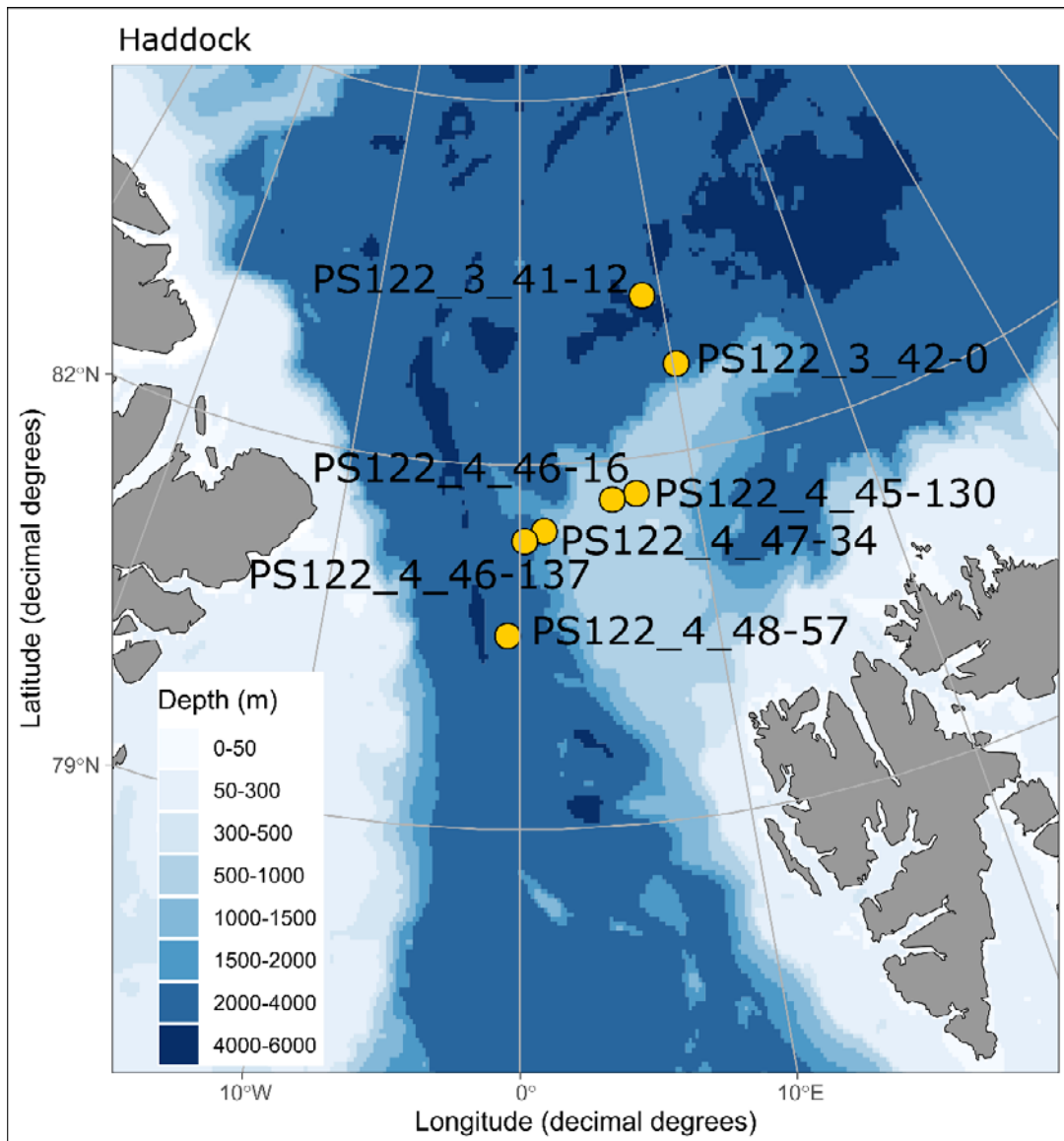
Figure 72: Map of the Arctic Ocean showing haddock (*Melanogrammus aeglefinus*) sampling sites (yellow bullets) in the the Atlantic Gateway during the MOSAiC expedition.

Table 41: Metadata of the 38 haddock (*Melanogrammus aeglefinus*) specimens collected in the Atlantic Gateway during the MOSAiC expedition genotyped at 98 SNPS on a MassARRAY iPLEX platform[160].

Sample code	Tissue code	Station	Latitude (°N)	Longitude(°E)	Marker
FR_10050	GN_0138	3_41-12	83.32	8.69	98 SNPS
FR_10051	GN_0143	3_41-12	83.32	8.69	98 SNPS
FR_10052	GN_0148	3_41-12	83.32	8.69	98 SNPS
FR_10048	GN_0128	3_41-12	83.32	8.69	98 SNPS
FR_10053	GN_0153	3_42_0	82.72	10.22	98 SNPS
FR_10057	GN_0173	3_42_0	82.72	10.22	98 SNPS
FR_10063	GN_0613	4_45-130	81.71	6.66	98 SNPS
FR_10065	GN_0615	4_45-130	81.71	6.66	98 SNPS
FR_10068	GN_0618	4_45-130	81.71	6.66	98 SNPS
FR_10067	GN_0617	4_45-130	81.71	6.66	98 SNPS
FR_10069	GN_0619	4_45-130	81.71	6.66	98 SNPS
FR_10062	GN_0612	4_45-130	81.71	6.66	98 SNPS
FR_10066	GN_0616	4_45-130	81.71	6.66	98 SNPS
FR_10064	GN_0614	4_45-130	81.71	6.66	98 SNPS
FR_10076	GN_0600	4_46-16	81.68	5.27	98 SNPS
FR_10077	GN_0601	4_46-16	81.68	5.27	98 SNPS
FR_10089	GN_0524	4_46-137	81.45	1.34	98 SNPS
FR_10088	GN_0526	4_46-137	81.45	1.34	98 SNPS
FR_10084	GN_0604	4_46-137	81.45	1.34	98 SNPS
FR_10093	GN_0607	4_46-137	81.45	1.34	98 SNPS
FR_10083	GN_0497	4_46-137	81.45	1.34	98 SNPS
FR_10082	GN_0495	4_46-137	81.45	1.34	98 SNPS
FR_10094	GN_0608	4_46-137	81.45	1.34	98 SNPS
FR_10092	GN_0605	4_46-137	81.45	1.34	98 SNPS
FR_10085	GN_0491	4_46-137	81.45	1.34	98 SNPS
FR_10086	GN_0493	4_46-137	81.45	1.34	98 SNPS
FR_10095	GN_0525	4_46-137	81.45	1.34	98 SNPS
FR_10091	GN_0494	4_46-137	81.45	1.34	98 SNPS
FR_10087	GN_0606	4_46-137	81.45	1.34	98 SNPS
FR_10090	GN_0498	4_46-137	81.45	1.34	98 SNPS
FR_10099	GN_0610	4_47-34	81.37	0.26	98 SNPS
FR_10098	GN_0609	4_47-34	81.37	0.26	98 SNPS
FR_10097	GN_0488	4_47-34	81.37	0.26	98 SNPS
FR_10112	GN_0602	4_48-57	80.59	-0.62	98 SNPS
FR_10108	GN_0489	4_48-57	80.59	-0.62	98 SNPS
FR_10109	GN_0492	4_48-57	80.59	-0.62	98 SNPS

---

[160] Berg PR, et al. (2021) Genetic structuring in Atlantic haddock contrasts with current management regimes. ICES Journal of Marine Science 78:1–13 [https://doi.org/10.1093/icesjms/fsaa204]

FR_10110	GN_0584	4_48-57	80.59	-0.62	98 SNPS

Polar cod (*Boreogadus saida*)

Polar cod, an ecological key species in the Arctic marine ecosystem, is widely distributed in both ice-covered and open waters of the Arctic Ocean and its adjacent seas, down to the Gulf of St. Lawrence on the North Atlantic coast and the Bering Sea in the North Pacific[161]. It is a key fish species in the biomass and energy flow from zooplankton to higher trophic levels of mammals and birds in the Arctic food web[162].

Twenty-one polar cod individuals from the CAO and Fram Strait were genotyped (**Figure 73**, **Table 42**)*.* The filtered data set used for population structure analysis of polar cod includes 615 SNPS. Genetic diversity is relatively low with observed and expected
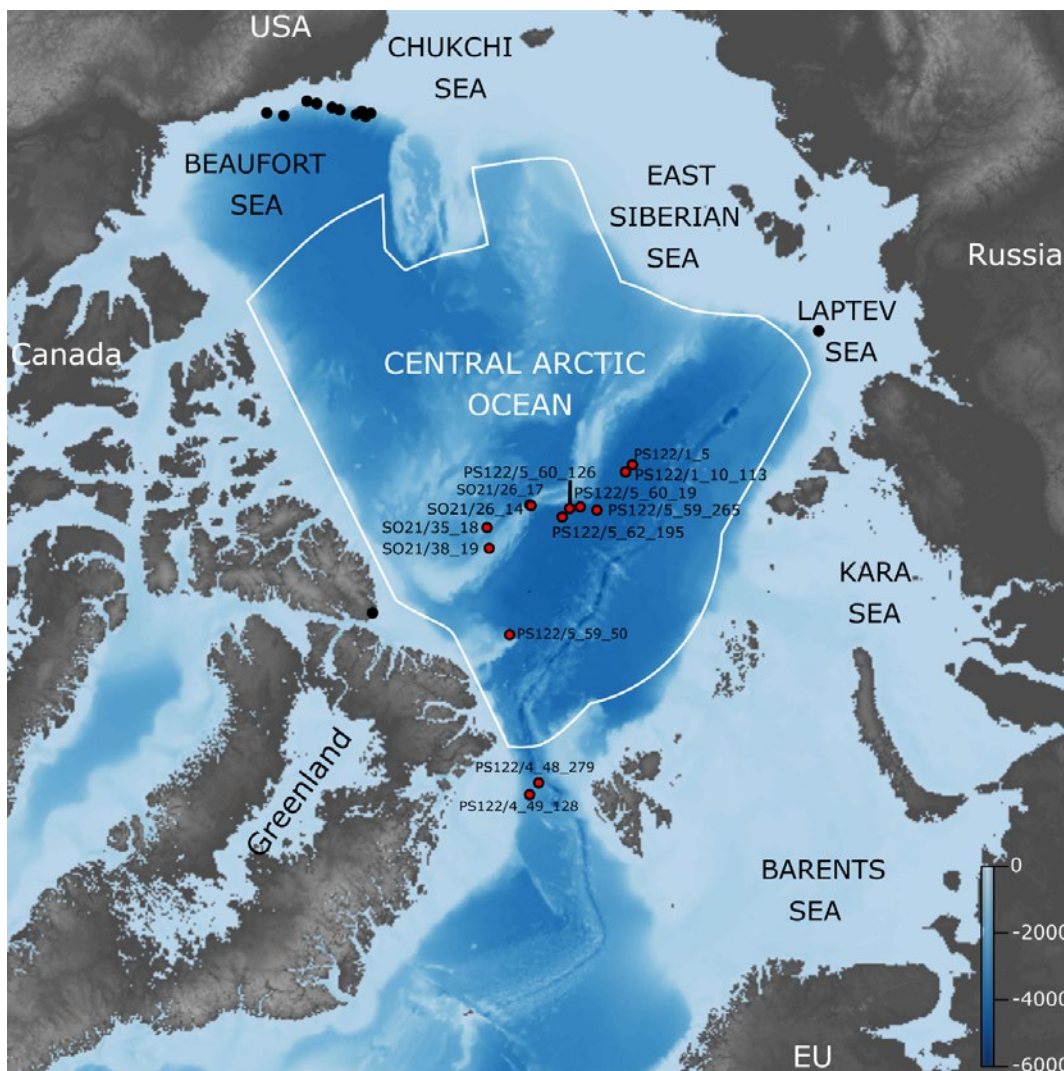
diversity (*Ho, He*) values of 0.20 and 0.21, respectively and inbreeding levels (F_{IS}) of 0.06.

---

[161] Marsh JM, Mueter FJ (2020) Influences of temperature, predators, and competitors on polar cod (*Boreogadus saida*) at the southern margin of their distribution. Polar Biology 43:995–1014 [https://doi.org/10.1007/s00300-019-02575-4]

[162] Dupont N, et al. (2020) Sea ice, temperature, and prey effects on annual variations in mean lengths of a key Arctic fish, *Boreogadus saida*, in the Barents Sea. ICES Journal of Marine Science 77:1796–1805 [https://doi.org/10.1093/icesjms/fsaa040]

Figure 73: Map of the Arctic Ocean showing polar cod *Boreogadus saida* sampling sites (red bullets) in the Central Arctic Ocean (CAO) and the Atlantic Gateway (AGAT) during the MOSAiC and SAS-Oden expeditions.

Table 42: Metadata of the 21 polar cod *Boreogadus saida* specimens collected in the Central Arctic Ocean (CAO) and the Atlantic Gateway (AGAT) during the MOSAiC and SAS-Oden expeditions, and genotyped at 615 SNPS on a Illumina Novaseq platform.

Sample code	Tissue code	Expedition	Station	Latitude (°N)	Longitude (°E)	Marker
FR_00005	GN_0077	MOSAiC	1_5	85.76	123.55	615 SNPS
FR_10021	GN_0002	MOSAiC	1_10-113	86.14	122.30	615 SNPS
FR_10014	GN_0013	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10015	GN_0021	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10016	GN_0025	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10017	GN_0029	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10019	GN_0033	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10022	GN_0037	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10023	GN_0042	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10118	GN_0046	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10119	GN_0050	MOSAiC	4_48-279	79.99	-0.65	615 SNPS
FR_10122	GN_0055	MOSAiC	4_49-128	79.52	-2.55	615 SNPS
FR_10125	GN_0057	MOSAiC	5_59-50	85.55	-16.22	615 SNPS
FR_10141	GN_0282	MOSAiC	5_59-265	87.78	104.85	615 SNPS
FR_10117	GN_0062	MOSAiC	5_60-19	88.34	114.96	615 SNPS
FR_10137	GN_0479	MOSAiC	5_60-126	88.74	120.36	615 SNPS
FR_10157	GN_0287	MOSAiC	5_62-195	89.14	110.83	615 SNPS
26-14_FISH_031	FISH_0629	SAS-Oden	26-14	89.14	-151.47	615 SNPS
26-100_FISH_028	FISH_001	SAS-Oden	26-17	89.10	-149.47	615 SNPS
35-100_FISH_069	FISH_026	SAS-Oden	35-18	87.87	-87.41	615 SNPS
38-100_FISH_060	FISH_002	SAS-Oden	38-19	87.77	-66.00	615 SNPS

A draft Principal Component Analysis (PCA) was performed with the 21 polar cod samples collected by the EFICA Consortium included in a large circum-Arctic data set (Verheye, Maes, Volckaert et al., unpublished). The majority of the samples collected by the EFICA Consortium from the CAO and the Atlantic Gateway clustered with samples collected in a large geographical area influenced by the Transpolar Drift, which spans the Laptev Sea to the North of Iceland and includes the European sector of the Arctic Ocean. A few of the samples from the CAO collected by the EFICA Consortium appeared genetically closer to samples coming from the Beaufort Sea, a region influenced by the Beaufort gyre. Young of the year are transported in surface waters over considerable distances from their spawning areas, while young polar cod migrate in association with sea ice. Hence, juvenile polar cod from the Siberian shelf might recruit to populations in the Greenland Sea, Canadian Archipelago, Denmark Strait, Svalbard and Barents Sea by advection with the Transpolar Drift[163,164,165].

Beaked redfish (*Sebastes mentella*)

---

[163] David C, et al. (2016) Under-ice distribution of polar cod *Boreogadus saida* in the central Arctic Ocean and their association with sea-ice habitat properties. Polar Biology 39:981-994 [https://doi.org/10.1007/s00300-015-1774-0]

[164] Maes S, et al. (2022) Comparative visual and DNA-based diet assessment extends the prey spectrum of polar cod *Boreogadus saida*. Marine Ecology Progress Series 698:139-154 https://doi.org/10.3354/meps14145

[165] Maes (2022) Population genomics and trophic ecology of polar cod (*Boreogadus saida*) in a changing Arctic Ocean. PhD thesis. KU Leuven, Leuven, Belgium [https://kuleuven.limo.libis.be/discovery/search?query=any,contains,LIRIAS3740414&tab=LIRIAS&search_scope=lirias_profile&vid=32KUL_KUL:Lirias&offset=0]

Beaked redfish is a common fish species in the North Atlantic Ocean, which is commercially exploited. It has a high degree of genetic population diversity and is differentiated in at least three ecotypes[166]. Tissue samples from four fish individuals were available (**Table 43**). However, after due consult with the SC07 project coordinator, it was decided not to genotype these four individuals because of the very small sample size and the lack of a reliable technique and published data for comparison.

Table 43: Metadata of the four beaked redfish (*Sebastes mentella*) specimens collected in the Atlantic Gateway during the MOSAiC expedition.

Sample code	Tissue code	Station	Latitude (°N)	Longitude (°E)
FR_10030	GN_0133	3_41-12	83.32	8.69
FR_10049	GN_0168	3_42_0	82.72	10.22
FR_10056	GN_0620	4_45-130	81.71	6.66
FR_10070	GN_0499	4_47-34	81.37	0.26

Dispersal routes of fish species in the Arctic Ocean

The Atlantic cod *Gadus morhua* specimens sampled in the Arctic Ocean probably originated from the North-East Atlantic cod (NEAC) stock, which is distributed from Lofoten to the western fjords of Svalbard and the Barents Sea. Temperature reconstructions suggest that the specimens sampled in the CAO and the Yermak Plateau area originated from both warm waters of the southern part in the NEAC distribution range (>5 °C) and cold waters in the northern Barents Sea or Svalbard fjords (-1 to 2 °C) (**Figure 70**, **Figure 74**).

The haddock *Melanogrammus aeglefinus* specimens sampled in the Yermak Plateau area probably belonged to the European Arctic (EA) stock, which has a similar distribution range as NEAC. Most of these haddock specimens originated from warmer waters in the southern part of the EA distribution range (>5 °C), but a few originated from colder waters as found in the southern Barents Sea (2-5 °C) (**Figure 70**, **Figure 74**).

The polar cod *Boreogadus saida* specimens sampled in the in the CAO belonged to two genetic clusters, one linked to the Transpolar Drift, and the second one linked to the Beaufort Gyre. These observations lead to the hypothesis that advection with the Transpolar Drift feeds all the way into the Icelandic stock in the Denmark Strait (**Figure 70**, **Figure 74**).

---

[166] Saha A, et al. (2017) Geographic extent of introgression in *Sebastes mentella* and its effect on genetic population structure. Evolutionary Applications 10:77-90 [https://doi.org/10.1111/eva.12429]
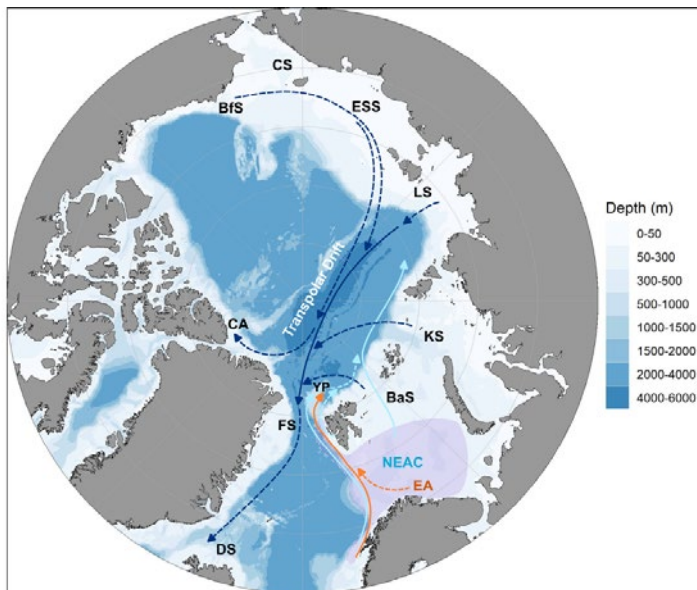
Figure 74: Possible dispersal routes of polar cod (dark-blue arrows), Atlantic cod (light-blue arrows), and haddock (orange arrows) in the Arctic Ocean, based on a combination of preliminary population genetics and life-history temperature reconstructions. The purple-shaded area indicates the approximate distribution range of the North-East Atlantic cod (NEAC) stock, which has a similar distribution as the European Arctic (EA) stock of haddock. Solid arrows indicate main dispersal routes, dashed arrows indicate additional routes from source populations to sink populations. BaS = Barents Sea, BfS = Beaufort Sea, DS = Denmark Strait, ESS = East Siberian Sea, FS = Fram Strait, KS = Kara Sea, LS = Laptev Sea.

5.6. Answers to the WP5 research questions

(1) Which species of fish occur in the CAO?

The samples analysed in WP5 demonstrate the occurrence of polar cod, Atlantic cod, ice cod, and black seasnail in the High Seas of the CAO. In addition, haddock and beaked redfish were present on the "doorstep" to the CAO, i.e., in the Atlantic inflow region on and near the Yermak Plateau and in the Fram Strait. Altogether, 24 fish species have either been reported to occur in the CAO or have been suggested to potentially expand into the CAO based on their life cycle and eco-physiological characteristics[10]. From the samples analysed in WP5, only polar cod, ice cod and black seasnail are included in this list. However, it cannot be concluded that any of the other 21 species were present or absent in the sampling area. Many of the 24 species are benthic fish, which were not targeted in the SC07 project. For some pelagic species that were not captured by longlines or gillnets it is likely that their abundance was too low or the fishing method inappropriate. The unexpected new observation of a large predatory fish species (Atlantic cod) in the CAO shows that scientific knowledge on the CAO ecosystem may still be incomplete with regard to nekton, and that the pelagic food web is probably more complex than previously thought.

(2) What is the origin of the fish in the CAO?

The combined results from temperature reconstructions and preliminary population genetic analyses indicate that Atlantic fish species enter the CAO from adjacent European shelf seas. Atlantic cod and haddock originate from stocks in the northern Norwegian Sea and the Barents Sea, including the Svalbard archipelago. Atlantic cod most likely belongs to the NEAC stock managed in ICES area 1, 2a and 2b, haddock belongs to the Northeast Arctic stock managed in ICES area 1 and 2b. The Arctic-endemic polar cod sampled in the CAO originate from spawning populations upstream of the Transpolar Drift. However, a small but significant proportion of the CAO population is probably advected with the Beaufort Gyre from as far west as the Beaufort Sea. The Transpolar Drift acts as a giant conveyor belt for polar cod, advecting fish from the Siberian shelf seas and even the American Arctic to the Canadian Archipelago and far south to the Denmark Strait.

(3) What is the diet, the condition and the nutritional quality of the fish?

Diet analysis of Atlantic cod, haddock and beaked redfish indicated that all fish were actively feeding, both in the Atlantic inflow region and in the CAO. Their diet consisted predominantly of macrozooplankton. In the shallower Yermak Plateau area, also benthic prey was taken. The overall contribution of fish and squid to the diet appeared low, but could be significant in terms of biomass. Both in the Yermak Plateau area and in the CAO, the analysis of condition index, gonadosomatic index and hepatosomatic index indicates that the majority of fish were in good health and had a nutritional value comparable to commercially harvested fish of the same species in the north Atlantic.

(4) What are the trophic interactions between zooplankton prey species and fish? (collaboration with WP4)

Results from ^{13}C and ^{15}N stable isotope analyses of fish muscle tissue confirmed that polar cod were dependent on ice-associated zooplankton such as *Apherusa glacialis* and *Calanus glacialis*. This impression was confirmed by the fatty acid pattern, in which diatom-associated fatty acids, indicative of ice algae, dominated over flagellate-dominated fatty acids, indicative of phytoplankton. In Atlantic cod, haddock, beaked redfish and ice cod, ^{13}C and ^{15}N stable isotope values clustered apart from polar cod and their zooplankton prey in the Arctic Ocean. This suggests that the time spent in the Arctic Ocean may not have been sufficient to transfer the isotopic signal of the Arctic pelagic food web to their muscle tissue, due to long turnover times. The relative marker fatty acid contributions showed that polar cod and black seasnail were predominantly associated with fatty acids indicative of diatoms. Haddock, Atlantic cod and beaked redfish caught in the Atlantic inflow region were associated with a stronger signal from pelagic flagellate fatty acids. Atlantic cod caught in the CAO showed a higher contribution of diatom-associated markers and low δ^{13}C values in relation to δ^{15}N levels, indicating a partial contribution of the cold-adapted food web of the CAO.

5.7. Relevance of the WP5 data for fish stock modelling

During the EFICA Workshop in preparation for the MOSAiC and SAS-Oden expeditions in June 2019[167], three potential categories of assessment methods, were identified depending on the data collected during surveys:
1. Data-rich stocks for quantitative assessments
2. Data-poor stocks for qualitative assessments
3. Data-poor stocks to indicate trends

Category 1 requires large data sets from repeated trawl surveys, which were logistically impossible to achieve during the MOSAiC and SAS-Oden expeditions, as already noted during the EFICA workshop.

For Category 2 a potential stock assessment model was identified (S6, Kokkalis et al. 2015)[168] requiring only size data. However, within SC07 only 53 gadoids caught with longlines or fishing rods, 22 polar cod caught with hand nets and traps, and one black seasnail caught with a zooplankton net, were available for size and mass measurements. Besides insufficient sample size, the applied sampling methods are highly size-selective and therefore the data do not provide a representative size distribution of the fish

[167] EFICA Consortium (2019) EFICA Workshop in preparation for the Polarstern and Oden expeditions for 2019 and 2020. Report to EASME/DG MARE. 32 pp. (EU internal document)
[168] Kokkalis A, et al. (2015) Limits to the reliability of size-based fishing status estimation for data-poor stocks. Fisheries Research 171:4–11 [https://doi.org/10.1016/j.fishres.2014.10.007]

populations. Furthermore, population genetic analyses showed that the Atlantic cod and haddock sampled in the Arctic Ocean belonged to stocks reproducing on the European shelf, which complicates assessment methods based on sampling in the Arctic Ocean alone. Therefore, the size and mass data analysed within WP5 are not suitable for stock modelling as required by the S6 model at this stage.

For Category 3 stock assessment and modelling approaches, high-quality habitat data are required along with genetic samples of the fish species that were analysed within WP5. The extensive ecosystem data of the MOSAiC and SAS2021 Oden expeditions would be more than suitable for this approach. The sample size of the fish collected by the EFICA Consortium is too small to estimate population parameters based on genetics[169], but when including a comprehensive circum-Arctic data set of polar cod population genetics (Verheye, Maes, Volckaert et al., unpublished), an assessment of the effective population size is feasible[170]. However, drawing on the extensive data on habitat properties and prey distribution collected during the MOSAiC and SAS2021 Oden expeditions, a first indication of the potential population size of predatory fish species in the CAO may be derived by modelling the spatial distribution of the carrying capacity (i.e., the rate of prey biomass production in relation to food demand of the fish populations[171] and niche modelling with stage-structured consumers[172].

5.8. Recommendations from WP5 for the JPSRM of the CAOFA

Design a standard "JPSRM longline". Longline fishing has proven to be a reliable tool. It should be used only when targets with strong backscatter are observed on the ship's echosounder. This method targets large predatory fish such as Atlantic cod, which could otherwise not be obtained. Perhaps lines can be designed to target smaller fish as well, but during the MOSAiC and SAS-Oden expeditions acoustic observations showed that especially the smaller fish fled as soon as something was lowered in the water column.

Deploy a standard mid-water trawl. This can be used ad-hoc when open water is available[173]. Patches of open water and even broader leads can occur between ice-floes due to wind forcing in the CAO, but their occurrence is unpredictable and often of short duration. Examples of equipment that can be used are Tucker trawl, RMT (Rectangular Mid-water Trawl) and IKMT (Isaacs-Kidd Midwater Trawl Net). This method targets mesopelagic macrozooplankton and smaller fish (polar cod, myctophids).

Deploy a Surface and Under-Ice Trawl (SUIT). This trawl was designed for under-ice fish (juvenile polar cod) and ice invertebrates[174].

Use the established stock assessment parameters.

---

[169] Ovenden, et al. (2016) Can estimates of genetic effective population size contribute to fisheries stock assessments? Journal of Fish Biology 89: 2505-2518 [https://doi.org/10.1111/jfb.13129]

[170] Nadachowska-Brzyska K, et al. (2022) Navigating the temporal continuum of effective population size. Methods in Ecology and Evolution 13:22–41 [https://doi.org/10.1111/2041-210X.13740]

[171] Myers RA (2001) Stock and recruitment: Generalizations about maximum reproductive rate, density dependence, and variability using meta-analytic approaches. ICES Journal of Marine Science 58:937–951 [https://doi.org/10.1006/jmsc.2001.1109]

[172] Nonaka E, Kuparinen A (2021) A modified niche model for generating food webs with stage-structured consumers: The stabilizing effects of life-history stages on complex food webs. Ecology and Evolution 11: 4101–4125 [https://doi.org/10.1002/ece3.7309]

[173] Ingvaldsen RB, et al. (2023) Under-ice observations by trawls and multi-frequency acoustics in the Central Arctic Ocean reveals abundance and composition of pelagic fauna. Scientific Reports 13:1000 [https://doi.org/10.1038/s41598-023-27957-x]

[174] Van Franeker JA, et al. (2012) The Surface and Under-Ice Trawl (SUIT). Technical Report https://www.researchgate.net/publication/297794282_The_Surface_and_Under_Ice_Trawl_SUIT#fullTextFileContent

Make a standard protocol for the JPSRM, including methodology. For example:
* Size and age distribution
* Fulton's condition index (K)
* gonadosomatic index (GSI)
* hepatosomatic index (HSI)
* Etc.

Additional recommended parameters to measure from fish samples.
Make a standard protocol for the JPSRM, including methodology. For example:
* Stomach (+ hindgut?) contents
* C:N ratio in muscle
* Stable isotopes (δ^{13}C, δ^{15}N) in muscle
* Fatty acids composition in muscle
* Otolith δ^{13}C for estimating field metabolic rate
* Otolith δ^{18}O for temperature reconstructions
* Standardised methods for population genetics (species-specific)

6. SEDIMENT OTOLITHS (WP6)

6.1. Research questions addressed by WP6

(1) Which fish species of have occurred in the CAO in the past as revealed by sediment otoliths?
(2) Do the age structure of the sediment otoliths reveal fish of all ages (i.e., not only the sympagic juveniles)?
(3) What was the ambient temperature of the fish when they were alive?

6.2. Data produced by WP6

Data file "EFICA_DATA_SC07-WP6"

For details of the Device Operations (date, time, geographical position, station depth), see files "MOSAiC_Device_operations" and "SAS-Oden_2021_Logbook"[175].

6.3. Human resources of WP6 and main responsibilities

Pauline Snoeijs-Leijonmalm (SU) coordination, ^{14}C dating; Julek Chawarski (SU) otolith identification, fish age, morphometrics; Flor Vermassen (SU) geological expertise, ^{14}C dating; Hauke Flores (AWI) δ^{13}C, δ^{18}O, ^{14}C dating; Kim Vane (AWI) otolith identification, fish age, δ^{13}C, δ^{18}O

Collaboration: The box-core sampling on-board the SAS-Oden 2021 expedition[176] was coordinated by Claudia Morys (SU) who was responsible for the SAS parameter "benthos". The otoliths in the stratified samples were collected from the benthos sub-sampling cores from which also the benthic macrofauna was sampled. The box-core sampling was also a collaboration with the SAS-Oden 2021 project on planktonic and sediment foraminifers (PI Helen Coxall, SU). Otoliths that possibly occur in the foraminifer project's subsamples from the box core samples are not included in the SC07 project. Sediment characteristics for the box core samples are analysed within the foraminifer project, and these data will be available at the Bolin Centre Database.

6.4. Methods used by WP6

Available samples

The otoliths elaborated within the SC07 project were collected from six sampling stations during the SAS-Oden expedition as described in detail in the Final Report from the field work in the CAO performed within SC06[177]. After microscopic inspection, 297 of the 305 "fish otoliths" collected from the box-core samples taken during the expedition were confirmed being fish otoliths. About half of the otoliths were collected from "stratified core sub-samples" taken with plexiglas corers that were pushed into the box core sample, i.e., these otoliths were collected from a known sediment stratum with 1 cm resolution. The other half of the otoliths were collected from unstratified "bulk samples", i.e., from the remaining sediment around the plexiglas cores in the box core sample (including sub-samples for other research projects).

---

[175] Bolin Centre Database [https://bolin.su.se/data/]
[176] Snoeijs-Leijonmalm, P. and the SAS-Oden 2021 Scientific Party (2022). Expedition Report SWEDARCTIC Synoptic Arctic Survey 2021 with icebreaker Oden. Swedish Polar Research Secretariat. 300 pp. [Link]
[177] Snoeijs-Leijonmalm P, et al. (2022) Ecosystem mapping in the Central Arctic Ocean (CAO) during the SAS-Oden expedition. Publications Office of the European Union [https://data.europa.eu/doi/10.2826/958629]

Sorting and identification of the otoliths

The sediment otoliths were in a varying state of preservation. Some were heavily encrusted by sediment or degraded by bioerosion from worms or by physical erosion from grinding by sediment grains. Since otolith shapes and morphology are characteristic for fish species, the most pristine otoliths that showed clear morphological traits were visually attributed to fish species in comparison with published images of fish otoliths from the Arctic, north Atlantic and north Pacific Oceans[178,179]. *Boreogadus* otoliths are more elongated in shape, smaller in size and weigh less than *Arctogadus* otoliths. *Arctogadus* otoliths are more rounded in shape, bulkier (thicker) and have deeper indentations on one end than *Boreogadus* otoliths[4]. The otoliths were visually assessed with regard to morphological features such as shape, ridges and lobes. Based on this assessment, the confidence in species identification was scored on an ordinal scale from 1 (low confidence) to 5 (high confidence).

Morphometric analyses

Morphometric analyses were performed on digital otolith images to support species identification. The morphometric data included major axis ("length"), minor axis ("width"), perimeter, and area, measured in mm using the software ImageJ[180]. In addition, the mass of each otolith was measured (in mg).

Fish age estimation

For age determination only well-preserved otoliths can be used, and for this purpose 86 sediment otoliths were selected. These 86 otoliths were embedded in epoxy resin to enable slicing through the core of the otolith. The otolith sections were then glued on a glass microscope slide, and a high-resolution photograph was taken under a Leica M205C stereo microscope with a Leica MC 170 HD camera. Age determination is commonly performed by assessing whole otoliths under transmitted light, but this procedure is more complicated for sediment otoliths than for otoliths from extant fish due to diagenetic alterations, such as autigenous calcite deposits and boreholes from worms. However, quality assurance of otolith aging was possible through comparisons with the AWI standard collection of polar cod otoliths taken from extant fish. The procedure included assessment of the otolith sections by two independent age readers. On the basis of the results a regression analysis was made with otolith weight as predictor variable and fish age as response variable. The fish age of the other otoliths was thereafter estimated from their mass using the regression model.

Radiocarbon dating

To reconstruct the time that had passed since an otolith was deposited, otoliths were analysed with standard radiocarbon (^{14}C) dating. The radioactive isotope ^{14}C is produced in the atmosphere, dissolves into the ocean and is incorporated in biogenic calcium carbonate, such as skeletons and otoliths of fish. Upon death of an organism, ^{14}C stops being incorporated in calcium carbonate in the bones, and ^{14}C slowly decreases over time due to radioactive decay with a half-life of 5700 ± 30 years. The time that has passed since the death of a fish can then be calculated from the ratio between stable and radioactive carbon isotopes. Pilot ^{14}C dating (5 samples) was performed at the at the

---

[178] Härkönen T (1986) Guide to the otoliths of the bony fishes of the northeast Atlantic. Danbiu Aps, Hellerup, Denmark. 256 pp.

[179] Campana SE (2004) Photographic atlas of fish otoliths of the northwest Atlantic Ocean. NRC Research Press, Ottowa, Ontario. 284 pp.

[180] https://imagej.nih.gov/ij/download.html

National Ocean Sciences AMS Facility (NOSAMS), Woods Hole Oceanographic Institution, USA. The rest of the ^{14}C analyses (36 otoliths) were performed at the AWI Radiocarbon Laboratory MICADAS (Mini Carbon Dating System)[181], following the standard operational procedure for carbonate samples (CHS)[182]. In short, the carbonate of the sample is transformed to CO_2 by acidic dissolution, and the CO_2 gas is measured in the MICADAS ^{14}C dating facility.

The scientific robustness of Arctic sediment age estimates based on ^{14}C dating is currently under debate. Recent research at AWI indicates that the precipitation of autigenous calcite and metal deposits (e.g., iron, manganese) on/in biogenic calcite such as bivalve shells or foraminifers lead to a significant overestimation of sediment age[183]. Microscopic investigation of the otoliths indicated that autigenous calcite deposits were present on the outer surface of many sediment otoliths, but not in the interior.

A method was developed aiming at minimising the potential error introduced by authigenic calcite and metal deposits[184] in ^{14}C dating of sediment otoliths. Each otolith was carefully cleaned to remove metal deposits before it was submitted for ^{14}C dating. In order to exclude the effect of authigenic calcite depositions on the age estimate, the ^{14}C dating was performed twice on each otolith sample. The first measurement was taken after ca. 30% of the outer sample mass containing the authigenic calcite deposits had been dissolved with 200 μL 0.05 M HCl, and secondly when the inner part was further hydrolysed with 200 μL H_3PO_4 for complete dissolution. The second measurement was considered not to be affected by authigenic calcite and represented the most reliable estimate of otolith residence time in the sediment. All values of the first measurement were higher than that of the second measurement for one and the same otolith. Only the second measurements are provided in this report.

Stable isotope analyses

The stable isotope ratios $\delta^{13}C$ and $\delta^{18}O$ in fish otoliths can be used to estimate field metabolic rate and ambient temperature, respectively, of a fish when it was alive[185,186]. All otoliths were cleaned as thoroughly as possible by slightly polishing or brushing the outside with ultra-pure water. After drying, the largest otoliths from stratified samples of the box corer stations were selected for taking a sample from their outer layer with a hand drill for $\delta^{13}C$ and $\delta^{18}O$ analyses at the Center for Marine Environmental Sciences (MARUM, Bremen) with the same method as used on extant fish in WP5 (**Figure 75**).

Most otoliths from stratified sediment samples were too small for micromilling the inner growth increments as performed in WP5 (**Figure 75**). Therefore, the selected sediment otoliths were not sectioned but a powder sample was taken from the outer surface of the whole otolith with a drill to target the most recent increment. The powder was collected into a glass vial for isotope analysis. The life-history timescale represented by an outer surface sample depends on otolith size and can be estimated by inspecting the increments of the respective otolith sections. For most small otoliths the sampled outer

[181] https://www.awi.de/en/science/geosciences/marine-geochemistry/micadas.html
[182] Mollenhauer G, et al. (2021) Standard operation procedures and performance of the MICADAS radiocarbon laboratory at Alfred Wegener Institute (AWI), Germany. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms 496:45-51 [https://doi.org/10.1016/j.nimb.2021.03.016]
[183] Wollenburg, JE (paper in review) Omnipresent authigenic calcite distorts Arctic radiocarbon chronology. Communications Earth and Environment
[184] Wollenburg JE, et al. (2023) Omnipresent authigenic calcite distorts Arctic radiocarbon chronology. Nature Communications Earth & Environment 4:136 [https://doi.org/10.1038/s43247-023-00802-9]
[185] Chung M, et al. (2019) Field metabolic rates of teleost fishes are recorded in otolith carbonate. Communications in Biology 2:1-10 [https://doi.org/10.1038/s42003-018-0266-5]
[186] Høie, H, et al. (2004) Temperature-dependent fractionation of stable oxygen isotopes in otoliths of juvenile cod (*Gadus morhua* L.). ICES Journal of Marine Science 61:243-251 [https://doi.org/10.1016/j.icesjms.2003.11.006]

surface layer represented an annual layer or ca. one life year of the individual, while this can be only a few months for the larger otoliths. This discrepancy originates from the size, where larger surfaces provide more material without going deeper into the structure. Such outer surface sampling allow for obtaining the most recent material, which is not possible when making sections as the outer layer is removed before micromilling due to the contact with resin that can influence isotope analysis.
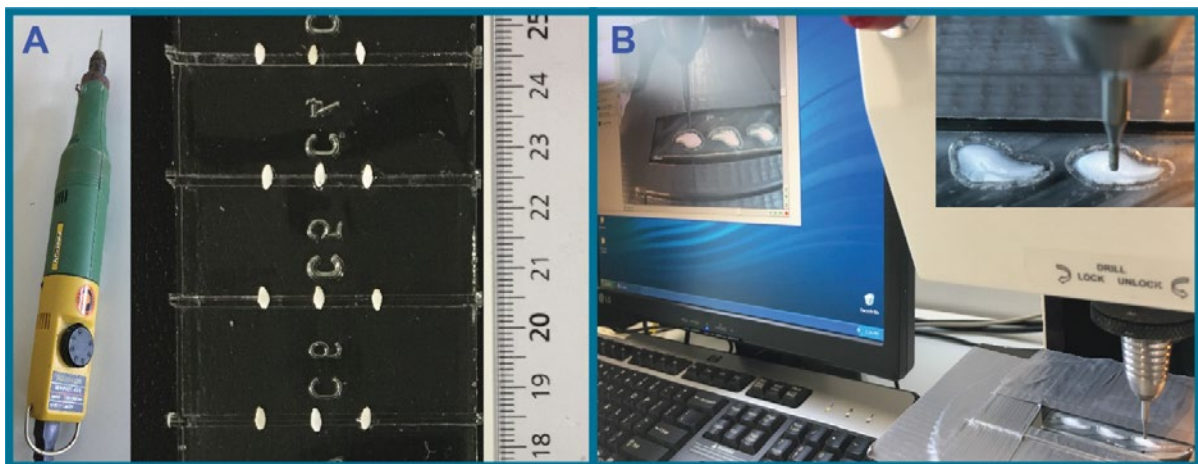
Figure 75: Methodology used for elaborating the otoliths collected during the MOSAiC expedition (otoliths of extant fish, WP5) and SAS-Oden expedition (sediment otoliths, WP6). (A) Sampling of the outer surface of whole otoliths fixed on a resin slate with a hand drill. The side of a drill with diamond coating was used to remove the outer surface of the otolith and the powder was collected into a glass vial for isotope analysis. (B) A computerised microdrill was used in WP5 to micromill targeted growth increments of each individual otolith. The procedure typically starts with micromilling a trench on the outer surface of the otolith section (enlarged inset photo) that effectively removes the outer carbonate layer of the otolith to prevent resin contamination in the sample for isotope analysis. Since micromilling of otolith sections is relatively coarse, this approach is mainly applied to larger otoliths. © Kim Vane

6.5. Results and discussion of WP6

Species identification

Due to heavy degradation of many of the otoliths, only 33% of the otoliths obtained a confidence score >3 for species identification. This illustrates that the species identifications presented in this report are associated with quite some uncertainty. Fifty-five of the 297 investigated otoliths were completely lacking distinctive morphological otolith features that are essential for species identification (**Table 44**). The majority of these unidentified otoliths had a gadoid-like shape. Of the remaining 242 otoliths, the majority (170) was identified as polar cod *Boreogadus saida*, 68 as ice cod *Arctogadus glacialis*, and four as *Paraliparis* spp., possibly representing two different species within this latter genus (**Table 44**). Although *Boreogadus* otoliths are photographically well-documented across life stages[187,188], *Arctogadus* otoliths are only documented for the adult stage. Juvenile otolith shapes of *Arctogadus* were extrapolated from the documented otolith shape development in other Gadidae species such as *Boreogadus saida* and *Gadus morhua*. Otolith shape development across life stages shows an overall

---

[187] Härkönen T (1986) Guide to the otoliths of the bony fishes of the northeast Atlantic. Danbiu Aps, Hellerup, Denmark. 256 pp.

[188] Campana SE (2004) Photographic atlas of fish otoliths of the northwest Atlantic Ocean. NRC Research Press, Ottowa, Ontario. 284 pp.

similar length-to-width ratio, although without pronounced morphological features such as lobes or indentures. *Arctogadus* otoliths are also distinct from *Boreogadus* by their higher weight relative to size. Thus, high-weight otoliths with similar length to width ratios but less pronounced features compared to *Arctogadus* adult otoliths were extrapolated to be representative of *Arctogadus* juvenile or young adult otoliths. *Spatial distribution of the sediment otoliths*

Of the 297 sediment otoliths, 139 were from stratified samples and 158 were from bulk samples (**Table 44**). An overwhelming majority of the otoliths in the stratified samples was found in the upper cm of the sediment, suggesting "recent" deposition on a geological time scale. Weak secondary peaks in otolith numbers at 5-8 cm and 12–14 cm of sediment depth might indicate variability in otolith deposition between different time periods (**Figure 76**). The highest number of polar cod otoliths was found at Station 38, and the lowest number at Station 60 (**Figure 77**). For ice cod, the highest number of otoliths was found at Station 50, and the lowest number at station 60 (**Figure 77**).

Table 44: Summary statistics of the 297 otoliths analysed.

Species	SAS-Oden 2021 Station						Total
	26	38	48	50	53	60	
Boreogadus saida	45	53	15	27	20	10	170
Arctogadus glacialis	4	21	4	28	10	1	68
Paraliparis spp.	2	1	1	-	-	-	4
Unidentified	24	11	9	8	1	2	55
Stratified	58	36	14	9	14	8	139
Bulk	17	50	15	54	17	5	158
Grand Total	75	86	29	63	31	13	297
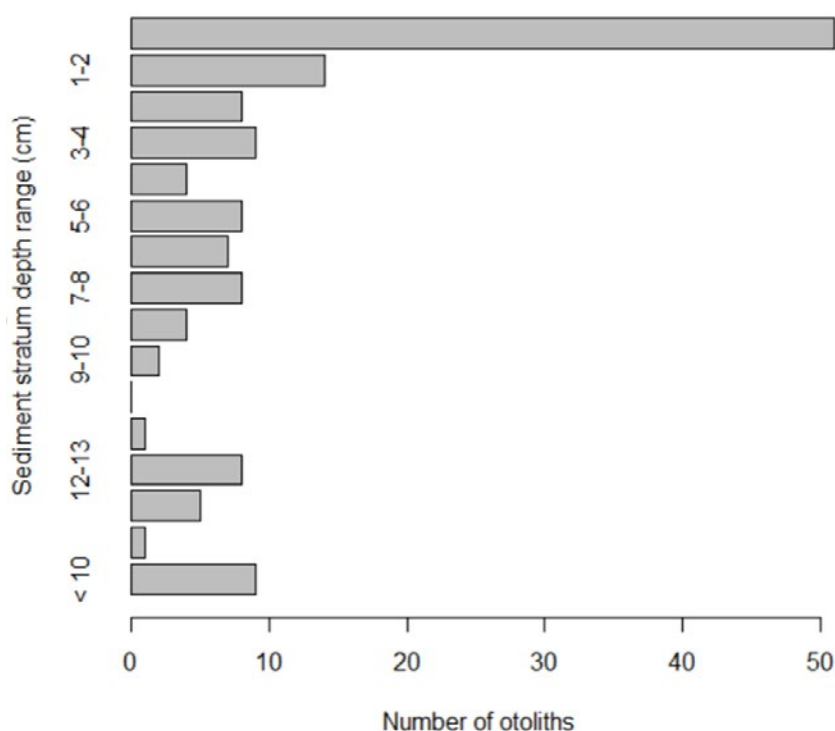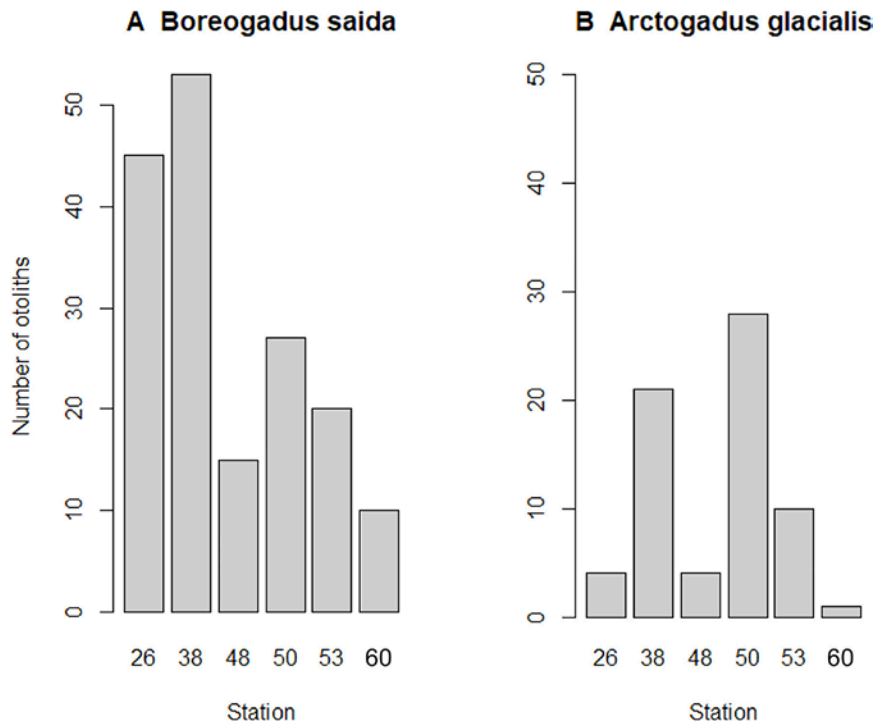
Figure 76: Depth distribution of 139 otoliths collected from stratified sediment samples during the SAS-Oden expedition. For each stratum of 1 cm the otoliths from all six sampling stations were combined. The stratum labelled "<10" indicates a pooled sample of all otoliths below 10 cm of sediment depth at Station 26 where the stratum 10-15 cm was combined.

Figure 77: Histograms of the number of otoliths collected per sampling station during the SAS-Oden 2021 expedition. (A) polar cod *Boreogadus saida*. (B) ice cod *Arctogadus glacialis*.

Otolith morphometrics

The shape characteristics of 241 of the 297 otoliths were in a condition allowing for morphometric measurements. The other 56 otoliths were degraded to an extent that could have compromised the accuracy of morphometric measurements. The morphometric analyses of all 241 suitable otoliths showed that the major axis ("length") of the sediment otoliths varied between 1.2 and 14.2 mm, the minor axis ("width) between 0.8 and 7.4 mm, the perimeter between 3.8 and 40.0 mm, and the area between 1.0 and 81.6 mm^2 (**Table 45**, **Figure 78**). The dimensions of ice cod *Arctogadus glacialis* otoliths were generally larger than those of polar cod *Boreogadus saida*, but there was an overlap in the otolith size of juveniles. The *Paraliparis* otoliths were the smallest.

Table 45: Summary statistics of morphometric parameters of the otoliths collected during the SAS-Oden 2021 expedition, excluding any otolith fragments or unidentifiable otoliths. n = number of otoliths analysed, Mean = arithmetic mean, Min = minimum value, Max = maximum value

Fish taxon	n	Major axis (mm)			Minor axis (mm)			Perimeter (mm)			Area (mm^2)		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max

	Major Axis [mm]			Minor Axis [mm]			Perimeter [mm]			Area [mm²]		
Boreogadus saida	150	4.6	2.2	8.4	1.9	1.0	3.1	11.6	5.7	21.4	7.4	1.9	20.4
Arctogadus glacialis	58	6.8	2.8	14.2	3.4	1.5	7.4	18.2	7.6	40.0	21.8	3.3	81.6
Paraliparis spp.	4	1.7	1.2	2.0	1.2	1.0	1.3	5.1	3.8	5.6	1.7	1.0	2.0
Unidentified	29	3.5	1.7	8.2	1.7	0.8	3.8	9.1	4.5	21.4	5.4	1.1	24.3
Total	241	5.0	1.2	14.2	2.2	0.8	7.4	12.8	3.8	40.0	10.5	1.0	81.6
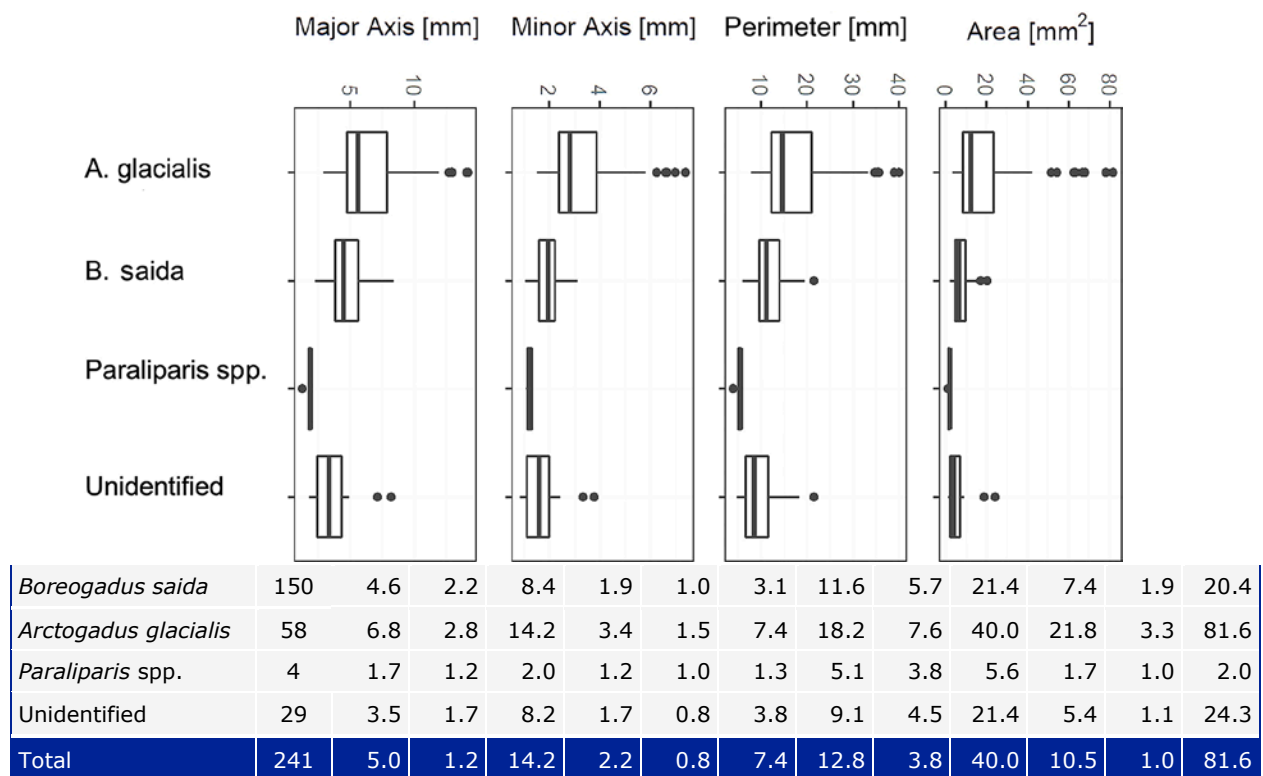
Figure 78: Boxplots summarising the morphometric parameters of the otoliths collected during the SAS-Oden 2021 expedition cf. Table 45.

Fish age

A robust age determination based on the consensus of two independent age readers was obtained for 32 out of 86 otoliths that had identifiable increments and no extensive boreholes. The majority of these 32 otoliths (22) was tentatively identified as *Boreogadus saida*. A correlation analysis of different morphometric parameters with age, as well as otolith mass, indicated that otolith mass was the best predictor of fish age for both *Boreogadus saida* and *Arctogadus glacialis*. A regression analysis yielded a significant linear relationship of log-transformed fish age with log-transformed otolith mass (p<0.001; adj. R^2=0.71). The regression equation obtained was:

Log (age) = 0.47684 * Log (otolith mass) - 0.21313

This relationship was used to reconstruct the tentative age distribution of polar cod and ice cod. The condition of the otoliths allowed reliable mass estimates in 163 polar cod and 66 ice cod otoliths. The reconstructed age data estimated that polar cod would have been between 1 and 5, and ice cod between 2 and 12 years old, which is in agreement with age data in the literature for these species. Most *Boreogadus saida* were 2-3 years old and most *Arctogadus glacialis* ca. 3-4 years (**Table 46**, **Figure 79**).

Table 46: Summary statistics of otolith mass, counts of age increments (annuli) and estimated age based on the otolith mass-age relationship estimated by the linear regression model for polar cod *Boreogadus saida* and ice cod *Arctogadus glacialis* from sediment otoliths sampled during the SAS-Oden expedition. n = number of otoliths, Mean = arithmetic mean, Min = minimum value, Max = maximum value

A Boreogadus saida

B Arctogadus glacialis
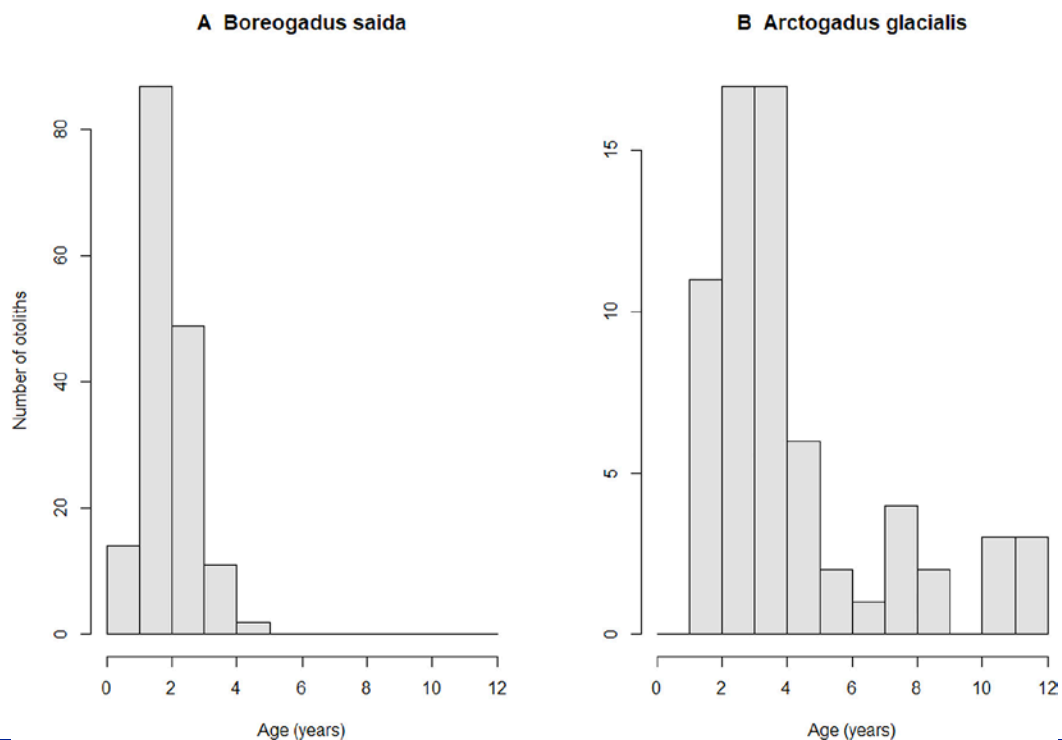
Fish taxon	Otolith mass (mg)				Annuli count				Estimated age			
	n	Mean	Min	Max	n	Mean	Min	Max	n	Mean	Min	Max
B. saida	163	7.7	0.6	40.2	22	2	1	3	163	2	1	5
A. glacialis	66	47.7	1.9	279.2	7	6	2	12	66	4	1	12
Unidentified	48	5.3	0.2	45.0	3	3	2	3	-	-	-	-
Total	281	16.6	0.2	279.2	32	3	1	12	229	2	1	12

Figure 79. Reconstructed age frequency distribution (A) polar cod *Boreogadus saida*. (B) ice cod *Arctogadus glacialis*.

Time of otolith deposition

As a pilot study, different materials (two foraminifer assemblages, two bivalve shells, one fish otolith) were ^{14}C dated to test whether the choice of material would impact ^{14}C dating results in years Before Present (yr BP) when the materials were deposited. The results showed a marked variability in deposition time in the surface layer, ranging from 8 660 (otolith) to 17 400 yr BP (bivalve), and with a foraminifer assemblage in-between (10 350 yr BP) (**Table 47**). In the deeper layer (9-10 cm), the obtained deposition times were very similar to each other – 30 700 (bivalve) and 29 800 yr BP (foraminifer assemblage) – but were much older than expected (pre-Holocene).

The results from this pilot study suggested that, to obtain the most accurate time for otolith deposition, it would be best to ^{14}C date the otoliths directly instead of inferring the time from bivalves or foraminifers deposited in the same sediment layer.

Table 47: Results of a pilot study for comparing ¹⁴C dating of various materials from two sediment strata in SAS-Oden box core SO21-53-Bx-11 (Station 53) analysed at the National Ocean Sciences AMS Facility (NOSAMS), Woods Hole Oceanographic Institution, USA. yr BP = years Before Present. Radiocarbon ages are given with one sigma confidence intervals.

Sediment stratum (cm)	Planktonic foraminifer assemblage (yr BP)	Bivalve (yr BP)	Otolith (yr BP)
0-1 cm	10 350 ± 55	17 400 ± 130	8 660 ± 40
9-10 cm	29 800 ± 710	30 700 ± 720	-

Table 48: Results of ¹⁴C dating of the inner fractions of 36 sediment otoliths taken from stratified sub-samples of five box-core samples during the SAS-Oden expedition and analysed at the MICADAS isotope facility (AWI). yr BP = years Before Present. Data for otoliths with deposition times <12 000 yr BP (Holocene) are coloured grey and data for otoliths with deposition times >12 000 yr BP (pre-Holocene) are coloured orange.

Fish species	Otolith ID	Station	Stratum (cm)	Otolith mass (mg)	Age at Death (yr)	yr BP	STD yr BP
B. saida	OTObc004	26	0-1	11.2	2.6	9 825	99
A. glacialis	OTObc046	26	0-1	50.6	5.0	2 144	65
B. saida	OTObc065	26	1-2	2.2	1.2	31 850	727
B. saida	OTObc069	26	2-3	9.2	2.3	3 507	104
B. saida	OTObc083	26	3-4	3.3	1.4	9 275	139
B. saida	OTObc030	26	4-5	12.0	2.0	32 425	759
B. saida	OTObc059	26	6-7	4.5	1.6	31 823	930
A. glacialis	OTObc094	26	7-8	34.3	5.0	35 257	1 131
B. saida	OTObc031	26	8-9	11.1	2.5	31 014	750
B. saida	OTObc016	26	9-10	1.6	1.0	42 488	1 195
B. saida	OTObc077	26	>10	3.1	1.4	42 046	1 146
A. glacialis	OTObc115	38	0-1	16.0	3.0	3 777	112
B. saida	OTObc128	38	0-1	2.1	1.2	41 650	1 196
A. glacialis	OTObc181	38	1-2	5.3	1.8	8 817	153
A. glacialis	OTObc200	38	2-3	2.6	1.3	8 907	94
A. glacialis	OTObc112	38	3-4	102.5	8.0	36 060	654
B. saida	OTObc107	38	4-5	3.7	1.5	31 774	750
B. saida	OTObc147	38	5-6	11.0	2.5	32 564	783
B. saida	OTObc127	38	6-7	10.1	2.4	35 558	1 005
B. saida	OTObc144	38	7-8	2.6	1.3	40 528	985
B. saida	OTObc145	38	12-13	4.3	1.6	43 321	1 325
B. saida	OTObc182	38	12-13	1.5	1.0	29 855	619
B. saida	OTObc187	38	13-14	1.9	1.1	45 421	1 653
B. saida	OTObc207	48	0-1	12.8	3.0	9 573	155
A. glacialis	OTObc212	48	0-1	238.3	12.0	17 438	253
B. saida	OTObc217	48	1-2	12.7	2.7	6 855	129
B. saida	OTObc258	48	3-4	13.1	2.8	9 678	152
B. saida	OTObc284	53	0-1	14.1	2.9	8 511	140
B. saida	OTObc292	53	0-1	13.5	2.8	9 381	91
B. saida	OTObc291	53	1-2	2.4	1.2	9 279	96
B. saida	OTObc268	53	3-4	2.6	1.3	5 446	77
B. saida	OTObc271	53	7-8	7.4	2.1	24 428	419
A. glacialis	OTObc353	53	12-13	20.9	3.4	28 464	559
A. glacialis	OTObc302	53	14-15	9.3	2.3	39 476	893
B. saida	OTObc285	60	0-1	10.3	3.0	-584	84
B. saida	OTObc378	60	13-14	12.8	2.7	5 946	78

The ^{14}C dating records of the inner otolith fractions showed an unexpectedly wide range. Deposition times ranged from approximately present-day in the surface sediment layer of Station 60 to ca. 45 000 years Before Present in an otolith from 13-14 cm sediment depth at Station 38 (**Table 48**). No otoliths deposited more than 12 000 years ago were expected because the Arctic Ocean was covered by a glacial ice sheet for about 150 000 years until 12 000 years ago, presumably allowing no biological production to support fish biomass. However, 20 out of the 36 otoliths were deposited more than 12 000 years ago (**Table 48**) if the ^{14}C dating is correct, which is assumed given the dating method developed in this project. Proof of living fish in the Arctic Ocean during the last ice age would be a scientific result with far-reaching implications in biological and geological science. However, before such a statement can be made with confidence, more samples must be analysed, and critical factors such as past sediment dynamics, ice sheet reconstructions and potential advection pathways from other regions must be carefully considered. The stratification is roughly reflected in the ^{14}C dating results (**Table 48**), but not always. For example, otoliths deposited more than 12 000 years ago were found in relatively shallow sediment at Stations 26, 38, and 48. This may be the result of sediment movements or even the sampling procedure with sub-sample coring.

To potentially identify more recently deposited otoliths, more otoliths need to be dated and the collection area in the CAO should be increased. This does not only imply new sampling because across geological research institutions there exist archived sediment cores from which otoliths can be retrieved and analysed. Such a wider collection could possibly also indicate the presence of more fish species in the CAO on short and long geological time scales. The area sampled during the SAS-Oden expedition is very central in the Arctic Ocean with a relatively consistent ice coverage and it can be concluded that two cold-adapted species, *Boreogadus saida* and *Arctogadus glacialis*, are dominant fish species here. Collecting sediment otoliths closer to the marginal ice zone (the transitional zone between open sea and dense drift ice) presents a higher chance of identifying additional fish species, including species that expand their distributions northward with climate change.

Stable isotopes

Of 86 sediment otoliths analysed, 75 yielded technically successful measurements of both δ^{13}C and δ^{18}O values. Of these, 51 otoliths were tentatively identified as *Boreogadus saida*, 14 as *Arctogadus glacialis*, and 10 were placed in the category "Unidentified" (**Table 49**). There was a significant difference in δ^{13}C values between the three groups (Kruskal-Wallis test, Chi^2=6.8, p<0.05), which was mainly driven by low δ^{13}C values (-3 to -5) of the 10 unidentified otoliths, whereas δ^{13}C values had a similar range (0.5 to -5.5) for *Boreogadus saida* and *Arctogadus glacialis* (**Figure 80**). δ^{18}O values did not significantly differ between the three groups (Kruskal-Wallis test, Chi^2=4.3, p>0.05).

Several caveats impede the calculation of scientifically robust temperature reconstructions at this point. Firstly, a major factor determining the reconstructed temperature in ancient otoliths is the assumed δ^{18}O value of the ambient water when the fish were living. The possibility to reconstruct accurate temperatures from δ^{18}O values in fish otoliths therefore depends on the correctness of the assumed δ^{18}O of water masses and their variability over time. In particular, when the Arctic Ocean was covered by a glacial ice sheet, the δ^{18}O values in the underlying water may have differed from present values. Secondly, the isotopic composition of the outer layer in sediment otoliths may have been influenced by chemical processes in the sediment, such as the precipitation of autigenous calcite, which could affect the isotopic composition of the outer surface of the

otolith[189]. Future robust temperature reconstructions will therefore depend on access to large datasets of δ¹⁸O values in waters of the Arctic Ocean and adjacent areas, their variability over geological time scales, and a higher number of δ¹⁸O measurements in sediment otolith samples unaffected by autigenous calcite.

Table 49: Stable isotope ratios δ¹³C and δ¹⁸O values expressed as per mille (‰) versus Vienna PeeDee Belemnite standard (VPDB) in sediment otoliths from the SAS-Oden 2021 expedition.

Fish species	Otolith ID	Station	Stratum (cm)	Otolith mass (mg)	Age at Death (yr)	yr BP	STD yr BP
A. glacialis	OTObc094	26	7-8	34.3	5.0	-3.62	3.08
A. glacialis	OTObc046	26	0-1	50.6	5.0	-0.22	3.76
A. glacialis	OTObc100	26	0-1	42.0	4.0	-2.17	3.21
A. glacialis	OTObc156	38	1-2	6.6	2.0	-5.12	3.25
A. glacialis	OTObc180	38	2-3	9.8	2.4	-3.55	2.27
A. glacialis	OTObc112	38	3-4	102.5	8.0	-2.36	4.07
A. glacialis	OTObc114	38	0-1	12.3	6.0	-2.85	1.75
A. glacialis	OTObc115	38	0-1	16.0	3.0	-3.28	2.79
A. glacialis	OTObc164	38	0-1	10.5	2.0	-4.57	4.10
A. glacialis	OTObc209	48	0-1	9.8	2.4	-3.95	3.12
A. glacialis	OTObc210	48	0-1	8.6	2.3	-3.68	2.37
A. glacialis	OTObc212	48	0-1	238.3	12.0	-2.37	5.83
A. glacialis	OTObc353	53	12-13	20.9	3.4	-3.11	1.47
A. glacialis	OTObc239	59	1-2	5.1	1.8	-3.81	2.62
B. saida	OTObc053	26	1-2	5.1	2.0	-3.29	0.38
B. saida	OTObc069	26	2-3	9.2	2.3	-2.35	0.14
B. saida	OTObc030	26	4-5	12.0	2.0	-3.52	2.23
B. saida	OTObc035	26	4-5	2.3	1.2	-2.56	3.01
B. saida	OTObc068	26	6-7	11.8	2.0	-1.87	2.98
B. saida	OTObc026	26	7-8	6.8	2.0	-4.77	2.88
B. saida	OTObc057	26	7-8	3.6	1.5	-5.02	3.60
B. saida	OTObc086	26	7-8	8.7	2.0	-2.30	2.05
B. saida	OTObc031	26	8-9	11.1	2.5	-2.71	2.74
B. saida	OTObc032	26	8-9	8.2	2.2	-1.88	1.60
B. saida	OTObc034	26	< 10	4.5	1.7	-3.97	2.76
B. saida	OTObc066	26	< 10	2.4	1.2	-4.49	3.11
B. saida	OTObc004	26	0-1	11.2	2.6	-3.62	2.32
B. saida	OTObc006	26	0-1	13.1	2.0	-3.45	4.21
B. saida	OTObc036	26	0-1	13.9	3.0	-3.19	3.24
B. saida	OTObc040	26	0-1	15.0	2.9	-1.87	1.56
B. saida	OTObc041	26	0-1	10.0	2.0	-3.83	1.90
B. saida	OTObc044	26	0-1	5.0	1.0	-3.10	3.11
B. saida	OTObc049	26	0-1	8.3	2.0	-3.17	1.71
B. saida	OTObc125	38	2-3	7.7	2.0	-3.20	2.23
B. saida	OTObc192	38	2-3	4.9	1.7	-4.19	-0.26
B. saida	OTObc136	38	4-5	2.8	1.3	-3.63	1.44
B. saida	OTObc147	38	5-6	11.0	2.5	-2.25	2.18
B. saida	OTObc185	38	5-6	5.1	1.8	-3.87	3.45
B. saida	OTObc127	38	6-7	10.1	2.4	-3.36	3.77
B. saida	OTObc101	38	11-12	5.2	1.8	-5.31	2.56
B. saida	OTObc138	38	12-13	11.2	2.6	-2.44	1.97
B. saida	OTObc176	38	12-13	11.6	2.0	-3.22	1.90
B. saida	OTObc189	38	12-13	5.0	1.7	-4.53	3.42

[189] Wollenburg JE, et al. (2023) Omnipresent authigenic calcite distorts Arctic radiocarbon chronology. Nature Communications Earth & Environment 4:136 [https://doi.org/10.1038/s43247-023-00802-9]

B. saida	OTObc113	38	0-1	2.5	1.3	-4.33	2.11
B. saida	OTObc119	38	0-1	11.9	2.0	-3.13	3.38
B. saida	OTObc135	38	0-1	No data	No data	-3.85	3.32
B. saida	OTObc177	38	0-1	5.9	2.0	-3.56	4.20
B. saida	OTObc199	38	0-1	3.9	1.0	-3.39	1.44
B. saida	OTObc217	48	1-2	12.7	2.7	-2.94	2.20
B. saida	OTObc258	48	3-4	13.1	2.8	-4.04	-0.03
B. saida	OTObc207	48	0-1	12.8	3.0	-2.70	-2.77
B. saida	OTObc211	48	0-1	11.0	3.0	-3.04	0.56
B. saida	OTObc293	50	3-4	24.3	3.7	-4.82	2.30
B. saida	OTObc290	53	1-2	9.1	2.0	-0.89	4.69
B. saida	OTObc234	53	0-1	7.7	2.1	-2.62	1.95
B. saida	OTObc279	53	0-1	7.7	2.1	-1.86	1.24
B. saida	OTObc281	53	0-1	10.7	3.0	-2.39	3.08
B. saida	OTObc284	53	0-1	14.1	2.9	-2.67	-0.09
B. saida	OTObc292	53	0-1	13.5	2.8	-3.27	2.08
B. saida	OTObc241	59	12-13	7.5	2.1	-3.08	3.29
B. saida	OTObc285	59	0-1	10.3	3.0	-4.89	4.06
B. saida	OTObc304	59	13-14	9.7	3.0	-2.48	1.98
B. saida	OTObc350	59	13-14	5.5	1.8	-3.71	1.76
B. saida	OTObc355	59	13-14	3.4	2.0	-4.42	2.99
B. saida	OTObc378	59	13-14	12.8	2.7	-3.90	4.21
Unidentified	OTObc018	26	1-2	2.5	2.0	-3.59	2.05
Unidentified	OTObc051	26	1-2	4.2	1.6	-4.17	1.43
Unidentified	OTObc097	26	6-7	6.9	2.0	-4.02	4.14
Unidentified	OTObc033	26	8-9	10.6	3.0	-3.93	3.42
Unidentified	OTObc052	26	0-1	9.7	2.4	-4.20	2.83
Unidentified	OTObc058	26	0-1	3.0	1.4	-3.96	2.51
Unidentified	OTObc162	38	2-3	3.2	1.4	-4.31	1.01
Unidentified	OTObc216	48	1-2	8.7	3.0	-4.36	3.02
Unidentified	OTObc367	53	9-10	1.5	1.0	-4.96	4.56
Unidentified	OTObc289	59	0-1	13.0	2.7	-2.91	3.61

Figure 80: Scatter plot of δ¹⁸O versus δ¹³C values from the outer surface of sediment otoliths of ice cod *Arctogadus glacialis*, polar cod *Boreogadus saida*, and unidentified otoliths cf. Table 49.

6.6. Answers to the WP6 research questions

(1) Which fish species of have occurred in the CAO in the past as revealed by sediment otoliths?

The dominance of polar cod *Boreogadus saida* and ice cod *Arctogadus glacialis* otoliths in the sediment samples provides a first indication that, both recently and in the past, no other abundant pelagic fish species would have occurred in the research area of the SAS-Oden expedition. The sediment otoliths provide evidence that *Boreogadus saida*, *Arctogadus glacialis*, and *Paraliparis* spp. have occurred in the CAO during the whole Holocene. The ¹⁴C dating results further suggest that these two species may even have been present in the Arctic Ocean during a period of the last ice age approximately 30 000 – 45 000 years before present. This time period roughly compares to published ¹⁴C dating results from (wrongly identified) Arctic sediment otoliths[190]. WP6 has used the latest knowledge on accurate dating of biogenic carbonate structures[191] by removing metal deposits, organic material/sediment and autigenous carbonate, and the ¹⁴C dating results reported here should be reliable.

(2) Do the age structure of the sediment otoliths reveal fish of all ages (i.e., not only the sympagic juveniles)?

[190] Hillaire-Marcel C, et al. (2022) Challenging radiocarbon chronostratigraphies in central Arctic Ocean sediment. Geophysical Research Letters 49:1-8 [https://doi.org/10.1029/2022GL100446]
[191] Wollenburg JE, et al. (2023) Omnipresent authigenic calcite distorts Arctic radiocarbon chronology. Nature Communications Earth & Environment 4:136 [https://doi.org/10.1038/s43247-023-00802-9]

Although otoliths from juvenile *Boreogadus saida* (1-2 years old) dominated in the sediment samples, many *Boreogadus saida* (up to 5 years) and *Arctogadus glacialis* (up to 12 years) were from adult fish as well. These WP6 results suggest that the overall distribution of these species in the CAO may even today contain a higher proportion of older individuals than previously assumed, which would support the hypothesis that juveniles use the CAO sea ice cover as habitat while adults contribute to the mesopelagic deep scattering layer (DSL) that was observed with hydroacoustics (WP2).

(3) What was the ambient temperature of the fish when they were alive?

The WP6 results show that otolith $\delta^{18}O$ values are generally in the same range as those measured in extant fish. This indicates that there is likely a high potential to reconstruct ambient temperatures when in-depth knowledge on the $\delta^{18}O$ values of Arctic water masses during the Holocene and the past ice age will become available. Hence, more measurements on otoliths from sediments and living fish from the CAO are recommended. A comprehensive database on present and past $\delta^{18}O$ values of Arctic waters will be necessary to provide scientifically robust reconstructed temperature estimates from otolith $\delta^{18}O$ values.

6.7. Relevance of the WP6 data for fish stock modelling

The WP6 results support that efforts of fish stock modelling should primarily focus on the two endemic Arctic gadoids *Boreogadus saida* and *Arctogadus glacialis* because the otoliths of these two species dominated the otolith assemblages in the deep-sea sediments of the CAO. The sediment otolith record shows that a higher proportion of the *Boreogadus saida* population in the CAO was older than what is known from data collected from under-ice polar cod[192], i.e., older than two years. This suggests that older fish of this species have been overlooked by the usually epipelagic sampling of *Boreogadus saida* in the CAO and should be considered in fish stock modelling. However, to obtain accurate data of the age and size structure of *Boreogadus saida* and *Arctogadus glacialis* populations in the CAO, it will be necessary to sample fish in the mesopelagic layer as otoliths are only an indirect measure.

The results of WP6 show that the sampling of sediment otoliths can provide important information about the potential (long-term) species composition, age and size structure of fish assemblages in areas where the sampling of living fish is impossible. Such knowledge can help to improve fish distribution models and niche modelling. Species identification of sediment otoliths can also aid in understanding the spatio-temporal changes in fish species composition in an area. Analyses of sediment otoliths thus can give an indication on how long a particular fish species has occurred in the CAO. When temperature reconstructions in sediment otoliths will be better constrained, these data can be valuable to construct temperature envelopes of fish in the CAO, which could be important for fish stock modelling and predictions of future changes in distribution ranges.

6.8. Recommendations from WP6 for the JPSRM of the CAOFA

Deep-sea sediment otoliths can provide useful data for the JPSRM. Otoliths indicate which species have dominated in a specific area in the past during periods with different environmental conditions, and can also show which species have invaded the area recently. When sampling of living fish is difficult, sediment otoliths can provide an

---

[192] Melnikov IA & Chernova NV (2013) Characteristics of under-ice swarming of polar cod *Boreogadus saida* (Gadidae) in the Central Arctic Ocean. Journal of Ichthyology 53:7-15 [https://doi.org/10.1134/s0032945213010086]

indication of which species to use in stock assessment and modelling. Sampling of deep-sea sediments can be performed during dedicated ecosystem expeditions to the CAO, but also use "ships of opportunity" with no or very limited biological sampling, e.g., geological surveys. Furthermore, the geological research institutions of the CAOFA parties likely host a wealth of sediment samples with otoliths that could be used to significantly extend the knowledge on past and present fish distributions in the Arctic Ocean.

7. ENVIRONMENTAL DNA (WP7)

7.1. Research questions addressed by WP7

(*1*) *What is the eDNA distribution of nekton (fish, squid) species in the CAO?*
(2) *What is the eDNA distribution of the major fish prey species in the CAO?*
(3) *What is the eDNA distribution of the major fish predator species in the CAO?*
(4) *How did the new bioinformatics pipelines work for the CAO?**

* WP7 included a pilot study to evaluate the use of eDNA data for answering the
   questions above through the development of relevant bioinformatics methods and
   resources.

7.2. Data produced by WP7

Data file "EFICA_DATA_SC07-WP7"

For details of the Device Operations (date, time, geographical position, station depth),
see files "MOSAiC_Device_operations" and "SAS-Oden_2021_Logbook"[193].

Storage of raw sequencing data: The MOSAiC molecular raw data are publicly accessible
at IMG/MER[194] under project number Gs0153906 in the search field, and this database is
still growing as more samples are still being processed by the US-DOE Joint Genome
Institute, USA. Preliminary results on the biodiversity variability in 16 of these samples
have been published[195]. The SAS-Oden molecular raw data will be archived and made
publicly accessible in the data repository hosted by the EMBL European Bioinformatics
Institute[196] in 2023.

In this report the MOSAiC and SAS-Oden results are compared with unpublished results
from previous expeditions with the Swedish icebreaker Oden to show the eDNA
distributions of particular species in a wider area, both inside and outside the CAO. These
expeditions are: "Arctic Ocean 2002" (AO02) in the outflow of Arctic water from the CAO
to the North Atlantic Ocean along the eastern coast of Greenland, "Beringia 2005" in the
Canada Basin, and "LomRog III 2012" in the CAO. The data owner of these data is
Pauline Snoeijs-Leijonmalm (SU) All three expeditions were carried out with the Swedish
icebreaker Oden and were organised by the Swedish Polar Research Secretariat (SPRS)

7.3. Human resources of WP7 and main responsibilities

Pauline Snoeijs-Leijonmalm (SU) coordination of the eDNA analyses and data analyses of
the results; Stefan Bertilsson (SLU) coordination of the laboratory and bioinformatics
analyses; Prune Leroy (SLU) coordination of the laboratory extractions and amplicon
sequencing; Marine VandeWalle, Javier Vargas Calle, Lauren Davies (SLU) DNA
extractions, QC, lab processing; Moritz Buck (SLU) bioinformatics of COI and 12S
amplicons and bioinformatics of fish genomic markers in metagenomic and
metatranscriptomic data sets; Allison Churcher (NBIS-UMU), John Sundh (NBIS-SU),
Marcin Kierczak (NBIS-UU) bioinformatics of fish genomes, COI and 12S markers in
metagenomic and metatranscriptomic data sets.

[193] Bolin Centre Database [https://bolin.su.se/data/]
[194] US-DOE Joint Genome Institute, USA [https://img.jgi.doe.gov/cgi-bin/mer/main.cgi]
[195] Mock T, et al. (2022). Multiomics in the central Arctic Ocean for benchmarking biodiversity change. PLoS
     Biology 20(10):e3001835 [https://doi.org/10.1371/journal.pbio.3001835]
[196] EMBL European Bioinformatics Institute [https://www.ebi.ac.uk]

Resources in addition to the EU SC07 funding: The development of new methodology in bioinformatics was made possible through an additional grant for bioinformatics support to existing projects (in this case the SC07 project) from the National Bioinformatics Infrastructure Sweden (NBIS). Supercomputing time was covered by an additional grant from the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at UPPMAX and the Swedish User and Project Repository (SUPR). The sequencing of the MOSAiC Core Samples was financed by a grant from the DOE Joint Genome Institute (JGI), Berkely, USA to a collective of MOSAiC scientists, including Pauline Snoeijs-Leijonmalm and Stefan Bertilsson (both EFICA Consortium partners).

7.4. Methods used by WP7

Available samples

For both he MOSAiC and the SAS-Oden expedition, eDNA samples from the water column were taken from a CTD rosette carrying 24 Niskin bottles of 12 L each (**Figure 81**). After retrieval of the Omics CTD (**Figure 82**), 16 pre-marked carboys were filled with 12 L of water each (from one Niskin bottle). In this way, the carboys were easy to carry and samples could not get mixed up. Omics water was sampled from four depths while the rosette was coming up: the chlorophyll maximum in the water column around 15-30 m ("ChlMax", measured by a fluorometer on the rosette on its way down), 100 m, the temperature maximum in the water column around 300 m ("TempMax", measured by the CTD on its way down), and 1000 m. The TempMax is in the Atlantic Water Layer where the DSL is found. If possible, Omics water was sampled from additional depths: 2000 m, 3000 m, and 10 m above the seafloor, if possible logistically and depending on water depth at the station.

eDNA samples from ice stations were taken from bulk water samples taken in blue carboys or as solid ice or snow. The sea ice was accessed from the ship either via the gangway or with a man-basket (**Figure 83**), and during MOSAiC by snow-mobile to avoid the artificial light from the ship in winter. On the ice, several potential sampling sites carefully investigated with ice thickness drills. This inspection was necessary to ensure safety, and to choose ice of an appropriate thickness according to scientific criteria and available time to sample ice cores. Other criteria to select sampling sites were accessibility of melt ponds and the potential to sample snow in summer (SAS-Oden). All ice station work followed the same design on both expeditions (**Figure 83**, **Figure 84**). A square 5 x 5 m coring field was identified. Cross-wise walk ways allowed access to the inner area of the coring field. The outer corners of the coring field were marked with red poles to ensure recognition in foggy conditions, especially when the floe needed to be temporarily abandoned for safety reasons.

For ice coring, teams of 2-5 people were working with 1-3 ice corers in parallel. Ice cores were sampled using 9-cm diameter Kovacs ice corers operated by battery-driven drilling engines. Each ice core was placed in a cradle and individually photographed before it was cut in appropriate pieces. Ice-habitat water (melt pond, brine, ice-seawater interface water) was sampled with a hand-operated membrane pump after the ice coring work was completed (**Figure 83 E**). The same pumps were used on both expeditions. Water from the ice-water interface was sampled through a borehole about 10 cm below the lower ice margin to avoid contamination by brine from the borehole. For sampling brine water, a new borehole was drilled in such a way that about 0.5 m of ice remained at the bottom of the borehole. This hole immediately filled up with brine water. Melt pond water was sampled about 10 cm below the melt-pond surface. To confirm that the sampled water was not contaminated by water from other sources (e.g., from the seawater below), the salinity was constantly monitored during ice-habitat water sampling. Water and ice samples were transported back to the ship as soon as possible to avoid freezing of the

water samples. Snow was sampled close to the coring field by shovelling the rather loose layer into six 20 L plastic buckets with a metal shovel (SAS-Oden).

Immediately after the field sampling, the water-column water from the CTD and the ice-habitat water (ice-seawater interface, brine, melt pond) was, as fast as possible, brought to laboratory containers on-board during both expeditions (**Figure 85**). These labs were unheated and illuminated with red light to not contaminate photosensitive cells by white light. In the labs, other scientists than those sampling from the CTD or on the ice had already prepared the peristaltic pumps with tubing and gamma-irradiated (sterile) sterile Sterivex™ pressure filter units with a polyethersulfone membrane of 0.22 µm pore size, which allows the filter to remain safely in the capsule during storage. This fast procedure was necessary for two reasons: (1) RNA (gene expression) changes very fast when environmental conditions change, and (2) the water in the carboys partly freezes when carboys with water are left outside, which may destroy cells. The RNA and DNA samples were filtered and immediately flash-frozen in liquid nitrogen and later stored at -80 °C.

Figure 81: Preparations and deployment of the Omics CTD. (A) The blue 20-L carboys for sampling the Omics water are brought to the CTD container on the aft deck. (B) Carboys waiting for sampling the water. (C) The CTD is taken out of the container. (D) CTD prepared for deployment. (E) CTD ready for deployment. (F) CTD being deployed. (A) © Serdar Sakinan, (B,C,D,E,F) © Pauline Snoeijs-Leijonmalm

The viral DNA samples (SAS-Oden only) were – after iron chloride treatment – filtered from the filtrate of the RNA and DNA filters (still containing viruses) on a 142 mm diameter Omnipore PTFE membrane filter with pore size 1.0 µm filter on a membrane filter with pore size 0.8 µm as a support filter, immediately flash-frozen in liquid nitrogen and later stored at -80 °C.

The ice samples were melted in plastic bags (MOSAiC) or the ice and snow samples were melted in containers with a tap (SAS-Oden, Figure **85 A**). The melting process took 30-40 hours. From the melted ice and snow only DNA samples (no RNA samples) were taken because RNA (gene expression) changes very fast when environmental conditions change

and 36 hours of melting generates an enormous change of environmental conditions, so that analysing RNA would be unrealistic.

Figure 82: Retrieval of the Omics CTD. © Yannis Arck

Table 50: Summary of Omics ice-habitat sampling during the SAS-Oden expedition. During the MOSAiC expedition ice sections were always short (10 or 20 cm), resulting in more samples per ice core of equal length.

Standard ice habitat	Sample description
Bottom ice	Lower 10 cm of 16 ice cores combined into one bulk sample
Centre-bottom ice	Lower half ice core minus 10 cm bottom ice of 8 ice cores combined into one bulk sample
Centre-top ice	Upper half ice core minus 10 cm top ice of 8 ice cores combined into one bulk sample
Top ice	Upper 10 cm of 16 ice cores combined into one bulk sample
Ice-seawater interface water	Pumped 40 L of water from the ice-seawater interface from boreholes through the ice into two 20-L carboys
Interstitial brackish brine water	Pumped 40 L of brine water from boreholes until 0.5 m above the ice-seawater interface into two 20-L carboys
Melt pond water	Pumped 40 L of melt pond water from 10 cm under the melt pond surface into two 20-L carboys
Snow	120 L of snow taken with a clean shovel into six clean 20-L buckets with lids

To the ice samples an appropriate amount of 0.22 µm pre-filtered seawater was added (50 mL water per cm ice of 9 cm in diameter) to prevent cells from bursting and the ice left to melt in the laboratory. The snow from the sampling buckets was combined in a 115-L container with a tap (SAS-Oden), but no seawater was added because the organisms in the snow are adapted to freshwater conditions.

Figure 83: Ice stations during the SAS-Oden expedition. (A) Scientists and polar bear guards on their way to sampling site close to the ship. (B) A crane-operated basket for bringing people from the ship to the ice and back. (C) Overview of an ice station very close to the ship. (D) Ice-coring. (E) Sampling water from ice habitats, i.e., water from melt ponds, brine, and sub-ice seawater. (A) © Hans-Jørgen Hansen, (B) © Julia Muchowski, (C) © Johan Wikner, (D) © Hauke Flores, (E) © Flor Vermassen

The volume of water filtered on one Sterivex™ filter varied between ca. 3 and 12 L depending on the "fullness" of the filter. A filter was full when the water started to drip slowly from the lower end of the filter and the filtration was stopped. Filter fullness depended on the number of particles or "sliminess" of the microbes in the water. Trying

to filter more water on a full filter would result in burst of the tubing of the peristaltic pump.

Figure 84: Ice coring during the SAS-Oden expedition. (A) A 5 x 5 m coring site has been selected and ice coring has started. (B) Measuring ice thickness through a coring hole.
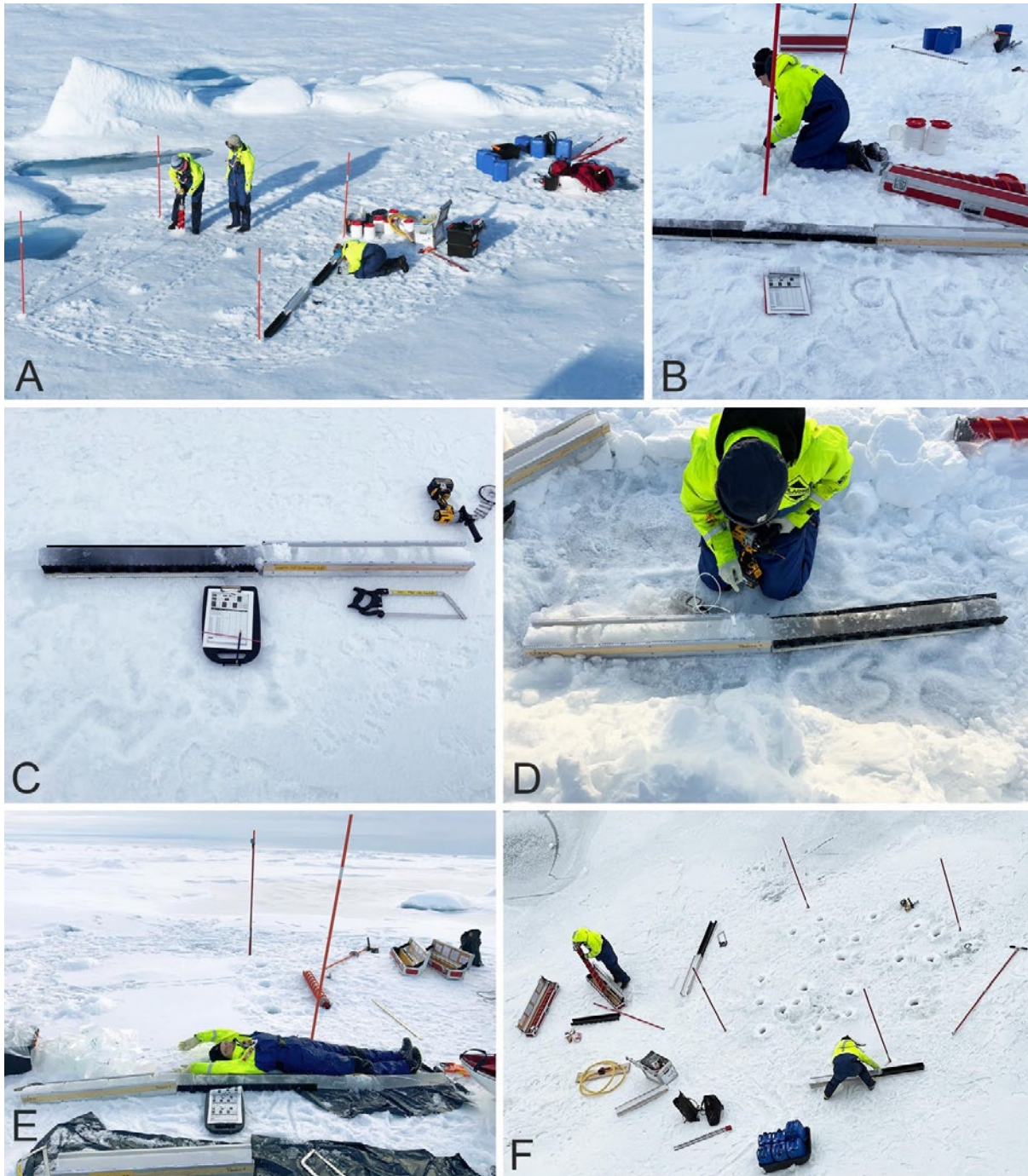
(C) Ice core in cradle with notebook, ice saw and drill. (D) Measuring a temperature profile in an ice core. (E) The longest ice core taken during the SAS-Oden expedition at Station 42, (F) Coring field after the coring had been completed. (A) © Johan Wikner, (B,C,D,E) © SAS-Oden sea-ice team, (F) © Pauline Snoeijs-Leijonmalm

Figure 85: Omics filtrations in the red-light laboratory containers during the SAS-Oden expedition. (A) The 115-L containers for ice and snow. The snow was sampled in the white buckets and the top and bottom ice-core sections in the containers with red lids. Filtered seawater was available in the blue containers. (B) The 50-L containers with the ice-seawater interface, brine, and melt pond bulk samples. (C) collecting water from the bulk samples into 10-L bottles with scales to read the volume. (D) The peristaltic pumps used. (E) Filtration on Sterivex filters. (F) After the filtrations the peristaltic-pump tubing was rinsed with MilliQ water (A,B,C,D, F) © Pauline Snoeijs-Leijonmalm, (E) © Swedish Polar Research Secretariat (SPRS)

The 10 omics subsamples taken at the six box-core sampling stations during SAS-Oden consisted of sediment cores in white 50-cm long polyvinyl chloride (PVC) tubes with an inner diameter of 3.8 cm (**Figure 86**). The tubes had been sterilised with 96% ethanol before sampling. Immediately after sampling, the outside of the tubes was cleaned and they were frozen at -80 °C as complete cores. Sub-sectioning of ca. 20 cm of the cores into 1-cm sections for RNA and DNA analyses took place in the home laboratory after the expedition. No sediment samples were taken during MOSAiC.

Figure 86: Sub-sampling a box core sample for Omics sediment samples during the SAS-Oden expedition. (A) Ten replicate Omics samples were taken with 10 white 3.8-cm diameter and 50-cm long polyethylene cores from each box core and immediately closed with aluminium foil on top. Twelve of the larger transparent plexiglas cores were used for benthic macroafauna and otoliths (WP6), and two were used for other projects. (B) The Omics samples covered both the darker and lighter-coloured sediment layers (presumably the Holocene and interglacial layers) (A,B) © Pauline Snoeijs-Leijonmalm

Set-up of the eDNA study

The eDNA study set-up (**Figure 87**) was a combination of amplicon sequencing and metagenomic / metatranscriptomic sequencing followed by a bioinformatics pipeline that uses international databases for taxonomic annotation, complemented with sequences from the fish and zooplankton project samples (see WP4 and WP5) to support annotation. The amplicons used were mitochondrial cytochrome c oxidase subunit I (COI), which is commonly used as a target for identification of metazoan genetic diversity at the species level, and the mitochondrially encoded 12S ribosomal RNA (12S) which is commonly used for targeting fish and mammals. The following data were produced:

Fish genome data derived from metagenomic sequencing of the MOSAiC and SAS-Oden expeditions "Omics" samples, tentative abundance data (per 10^9 reads) for fish genome records.

COI-marker data (COI = mitochondrial cytochrome c oxidase subunit I) derived from metagenomic and metatranscriptomic sequencing of the MOSAiC and SAS-Oden

expeditions "Omics" samples, taxon name assigned by molecular taxonomic database, tentative abundance data (per 10^9 reads) for COI-marker records. s

COI-amplicon data (COI = mitochondrial cytochrome c oxidase subunit I) from PCR amplification of DNA extracts from the MOSAiC and SAS-Oden expeditions "Omics" samples, presence/ absence data (0 or 1) for COI-amplicon records.

12S-amplicon data (12S = mitochondrially encoded 12S ribosomal RNA) from PCR amplification of DNA extracts from the MOSAiC and SAS-Oden expeditions "Omics" samples, presence/ absence data (0 or 1) for 12S-amplicon records.
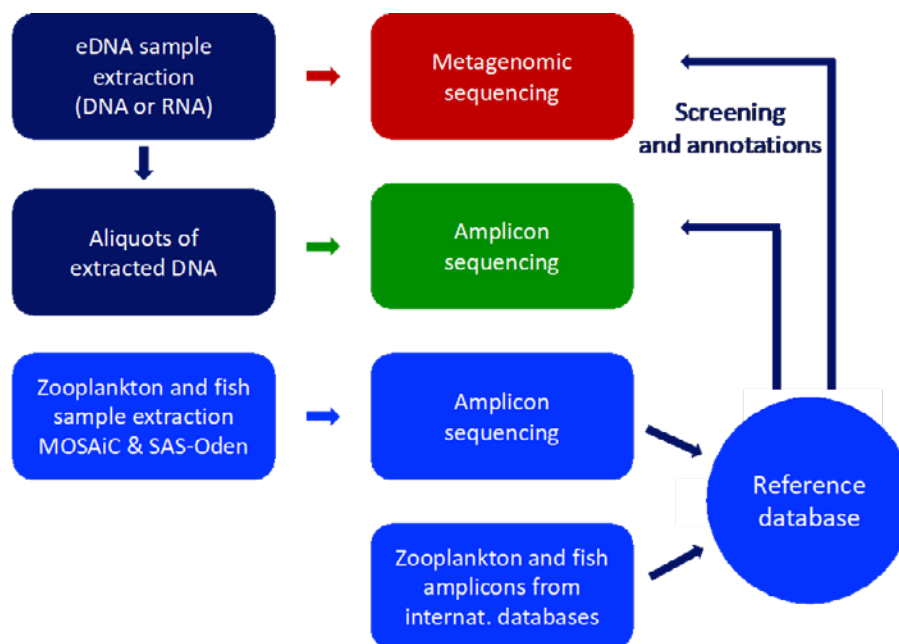
Figure 87: Schematic overview of the set-up of the eDNA study of the SC07 project. The amplicons selected were the COI and 12S markers. To the left the three types of extractions for metagenomic/metatranscriptomic sequencing, amplicon sequencing of water samples, and amplicon sequencing of individual faunal specimens, respectively. To the right the input to the database package from international databases as well as specimen sequencing and the screening and annotation pipelines.

Extractions

Types of extractions: Two types of extractions were performed: DNA extractions and RNA extractions. Different aliquots of the same DNA extraction were later used for metagenomic sequencing (no PCR amplification) and for amplicon sequencing (with PCR amplification). One aliquot of the RNA extraction was used for metatranscriptomic sequencing (no PCR amplification).

Laboratories involved: The MOSAiC Expedition Core DNA and RNA samples were extracted by the MOSAiC Omics Team partners at the AWI (not by the EFICA Consortium). The SAS-Oden Expedition Core Samples and the MOSAiC and SAS-Oden MIME samples were extracted at SLU (EFICA Consortium partner). For all samples the same extraction protocols were used at AWI and SLU for DNA and RNA.

Extraction procedures: A previously validated method based on Qiagen "PowerWater" kit according to the manufacturer's protocol for nucleic acid extractions from materials

captured on Sterivex™ filters were used to retrieve total eDNA. Prior to extraction the Sterivex™ filter capsules were broken open in a sterile laboratory environment to retrieve the filter for extraction because it was found that more DNA/RNA could be extracted compared to performing the extraction inside the filer capsule. RNA was extracted from an additional set of samples using Qiagen "RNeasy" mini kit according to the manufacturer's protocol with a first step of bead beating and an additional final step of DNaseI treatment. For the 50 samples collected within the "Arctic Virus" project[197], total nucleic acid (DNA+RNA) was extracted according to Promega's "Maxwell RSC Enviro Total Nucleic Acid" kit with several steps – a first step of ascorbic acid treatment to remove iron flocculates and a last step of RNase treatment followed by a DNA clean-up step with Qiagen "DNeasy PowerClean Pro Cleanup" kit. For the sediment samples, DNA was extracted according to "MP FastDNA Spin kit" for soil with an additional step of proteinase K treatment.

Extraction yield: The DNA and RNA concentrations obtained from the extractions were highly variable with the majority featuring sufficient concentrations for direct library preparation and shotgun sequencing (metagenomics and metatranscriptomics). A number of samples, mainly from deeper strata and winter during the MOSAiC expedition, and from deeper strata and sediments from the SAS-Oden expedition had too low nucleic acid concentrations (below 2 ng μL^{-1}) to enable shotgun sequencing. However, most of these samples could still be used for COI and 12S amplicon MiSeq sequencing (SLU Uppsala, in-house) after PCR amplification.

Extra reference samples: To enable species identification based on COI and 12S-amplicons, and as a complement to on-line reference databases, DNA from 77 zooplankton individuals (17 species, **Table 51**) and 30 fish individuals (6 species, **Table 52**) collected during the SAS-Oden and LomRogIII expeditions were extracted using "MN Nucleospin tissue kit" with a 2-hour proteinase K treatment, and in the case of species with hard exoskeletons (amphipods, copepods), a manual tissue disruption step using a pestle in a 1.5-mL Eppendorf tube was used.

Table 51: List of the 77 individuals of 17 zooplankton taxa caught during the SAS-Oden expedition, complemented with some individuals from other samples used for COI and 12S amplicon sequencing to complement the reference databases. LomRog III = Oden expedition to the CAO in 2012.

Fish species	Group	Expedition	Specimens in single-species sample	Specimens in mixed samples SAS-Oden
Cyclocaris guilelmi	Amphipod	SAS-Oden	2	3
Eusirus holmii	Amphipod	SAS-Oden	9	
Lanceola clausi	Amphipod	SAS-Oden	1	
Themisto abyssorum	Amphipod	SAS-Oden	3	
Themisto libellula	Amphipod	SAS-Oden	3	
Aetideopsis rostrata	Copepod	SAS-Oden		1
Calanus glacialis	Copepod	SAS-Oden + LomRog III	5	1
Calanus hyperboreus	Copepod	SAS-Oden + LomRog III	5	3
Metridia longa	Copepod	Lomrog III	4	2
Oithona similis	Copepod	SAS-Oden		1
Oncaea sp.	Copepod	SAS-Oden		1
Paraeuchaeta sp.	Copepod	Lomrog III	8	
Hymenodora gracilis	Decapod	SAS-Oden	3	
Botrynema ellinorae/brucei	Hydrozoan	SAS-Oden	11	2

[197] Snoeijs-Leijonmalm, P. and the SAS-Oden 2021 Scientific Party (2022). Expedition Report SWEDARCTIC Synoptic Arctic Survey 2021 with icebreaker Oden. Swedish Polar Research Secretariat. 300 pp. [[Link](#)]

Halocyprida_XXX	Ostracod	SAS-Oden		1
Eukrohnia hamata	Chaetognath	SAS-Oden		2
Pseudosagitta maxima	Chaetognath	SAS-Oden	4	2
Total			**58**	**19**

Table 52: List of the 30 individuals of six fish species caught during the MOSAiC and SAS-Oden expeditions used for COI and 12S amplicon sequencing to complement the reference databases. * = this individual was initially identified as *Arctogadus glacilis* and is currently being analysed further with regard to species identity.

Fish species	Expedition	Number of individuals
Boreogadus saida (polar cod)	SAS-Oden	3
Gadus chalcogrammus (walleye pollock) *	MOSAiC	1
Gadus morhua (Atlantic cod)	MOSAiC	14
Melanogrammus aeglefinus (haddock)	MOSAiC	10
Sebastes mentella (beaked redfish)	MOSAiC	1
Paraliparis bathybius (black seasnail)	MOSAiC	1

Sequencing

Metagenomes and metatranscriptomes: The MOSAiC Expedition Core Samples that were extracted by the MOSAiC Omics Team partners at the AWI were sequenced by the DOE Joint Genome Institute (JGI)[198,199]. The SAS-Oden Expedition Core Samples and the MOSAiC and SAS-Oden MIME samples that were extracted at SLU were sequenced at the Swedish National Genomics Infrastructure hosted by the Science for Life Laboratory (SLL)[200], Sweden. Shotgun sequencing was performed by multiplexed Illumina NovaSeq 6000 on 10 full S4-300 v 1.5 flowcells (350 bp insert size), aiming for at least 20 Gb of sequence data from each sample.

COI- and 12S-amplicons (PCR-based): Since several samples that were collected in winter or at deeper strata (deep water column and sediments) had too low DNA/RNA concentrations for direct sequencing, a PCR-based approach was used to target specific genetic marker. The sequence data from this approach was used to complement the direct shotgun sequencing data and were a valuable addition to the shotgun eDNA analyses. While these assays are not quantitative and may be subject to primer bias, they can still enable detection of low-density targets by enrichment through PCR and therefore fill gaps in the shotgun data set. The mitochondrial cytochrome c oxidase subunit I (COI) is commonly used as a target for identification of metazoan genetic diversity at the species level. After surveying previously published studies it was decided to amplify a short (313 bp) fragment of the highly variable mitochondrial COI region with previously validated primers (mlCOIintF[201] and jgHCO2198[202]) (**Table 53**). Primer jgHCO2198 is a modification of the previously commonly used primer HCO2198[203]. jgHCO2198 contains inosine and degenerate bases, allowing detection of most marine

---

[198] https://genome.jgi.doe.gov/portal

[199] Mock T, et al. (2022). Multiomics in the central Arctic Ocean for benchmarking biodiversity change. PLoS Biology 20(10):e3001835 [https://doi.org/10.1371/journal.pbio.3001835]

[200] https://www.scilifelab.se

[201] Leray M, et al. (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Frontiers in Zoology 10:34 [https://doi.org/10.1186/1742-9994-10-34]

[202] Geller J, et al. (2013) Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Molecular Ecology Resources 13:851-861 [https://doi.org/10.1111/1755-0998.12138]

[203] Folmer O, et al. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294-299. [https://www.mbari.org/wp-content/uploads/2016/01/Folmer_94MMBB.pdf]

invertebrate species, including zooplankton species that were missed by HCO2198. These modified primers were validated on DNA. To more specifically target and enrich genetic markers for fish, a second pair of primers matching the variable region of mitochondrial 12S rDNA was selected (**Table 53**). These primers, MiFish-U-F and MiFish-U-R were designed and validated by Miya et al.[204] and have been validated *in silico* (tested via computer simulations). These specific primers were validated on DNA extracted from 31 selected fish specimens corresponding to the six key species collected during the SAS-Oden expedition (see WP5). Also, the absence of amplification of 12S for the 77 zooplankton specimens was tested. These specific primers were validated on DNA extracted from 30 selected fish specimens corresponding to the six key species collected during the SAS-Oden expedition (see WP5). The specificity of the 12S rRNA assay was also tested, demonstrating that there was no amplification of 12S from any of the 77 zooplankton reference specimens. For CO1 and 12S amplicon sequencing DNA/RNA aliquots left after shotgun sequencing were used. In order to save material, a multiplex PCR approach was chosen in which both genetic markers were targeted and sequenced in the same reaction. COI and 12S reads were retrieved for most samples but not all. It was noticed that 12S primers also amplified bacterial 16S rRNA sequences as well due to high sequence similarity, which has been noticed previously by others[205]. While it was possible to filter out these contaminant sequences in the bioinformatics pipeline, this limited the number of true target sequences in many of the samples.

Table 53: List of the primers used for the COI and 12-S amplicon studies in the SC07 project.

Target region	Primer name	Reference
Mitochondrial DNA cytochrome c oxidase subunit I (COI)	mlCOIintF	Leray et al.[206]
Mitochondrial DNA cytochrome c oxidase subunit I (COI)	jgHCO2198	Geller et al.[207]
Mitochondrial 12S ribosomal DNA (rDNA)	MiFish-U-F	Miya et al.[208]
Mitochondrial 12S ribosomal DNA (rDNA)	MiFish-U-R	

Summary of the available data sets

The eDNA samples consisted of a mixture of microbial cells and environmental nucleic acids (eDNA) collected from water samples by membrane filtration through 0.22 μm pores (Sterivex™). Using the extraction and sequencing methods described below, four data sets that form the basis for the bioinformatic analyses presented in this report were built (**Table 54**). Successful shotgun sequencing data were obtained for 401 DNA samples and 455 RNA samples, and successful MiSeq amplicon sequencing data were obtained for 535 DNA samples.

---

[204] Miya M, et al. (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. Royal Society Open Science 2:150088 [https://doi.org/10.1098/rsos.150088]

[205] Gold Z, et al. (2020) FishCARD: Fish 12S California Current Species Reference Database for Enhanced Metabarcoding Efforts. [https://www.authorea.com/users/330161/articles/457029-fishcard-fish-12s-california-current-specific-reference-database-for-enhanced-metabarcoding-efforts]

[206] Leray M, et al. (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Frontiers in Zoology 10:34 [https://doi.org/10.1186/1742-9994-10-34]

[207] Geller J, et al. (2013) Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Molecular Ecology Resources 13:851-861 [https://doi.org/10.1111/1755-0998.12138]

[208] Miya M, et al. (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. Royal Society Open Science 2:150088 [https://doi.org/10.1098/rsos.150088]

Fish genome mapping derived from metagenomic sequencing

Fish diversity in the extreme cold and nutrient-poor CAO is very limited. All commercial fish species known to occur here have a draft genome available. However, no genomes were available for a group of non-commercial demersal fish species, such as eelpouts (Zoarcidae) and snailfish (Liparidae) previously reported from the CAO (**Table 55**). Unfortunately, no genome from the cephalopod *Gonatus fabricii* observed during MOSAiC was available.

Table 54: Summary of the available sequences from the MOSAiC and SAS-Oden expeditions for the analyses presented in this report. Metagenomes and metatranscriptomes were obtained by shotgun sequencing at the DOE Joint Genome Institute (JGI), USA and the Science for Life Lab (SLL), Sweden. Amplicon = sequencing obtained by DNA extraction, PCR amplification and MiSeq sequencing (SLU Uppsala, in-house).

Data set	MOSAiC ice AWI/JGI	MOSAiC water AWI/JGI	MOSAiC ice SLU/SLL	MOSAiC water SLU/SLL	SAS-Oden Ice SLU/SLL	SAS-Oden water SLU/SLL	SAS-Oden sediment SLU/SLL	Total
Metagenomes Nr of samples	33	35	13	21	156	127	16	401
Metagenomes Nr of reads	16.3×10^9	16.7×10^9	2.1×10^9	3.8×10^9	37.5×10^9	32.2×10^9	5.2×10^9	113.7×10^9
Metatranscriptomes Nr of samples *	12	5	174	252	4	8	16	455
Metatranscriptomes Nr of reads	2.4×10^9	0.8×10^9	21.8×10^9	31.4×10^9	0.5×10^9	1.1×10^9	-	57.9×10^9
Amplicon data set Nr of samples	-	-	118	60	164	127	66	535
Amplicon data set Nr of reads	-	-	1 744 372	1 022 761	2 329 880	1 536 650	780 879	7 414 542

A robust and sensitive mapping-driven approach was developed to catch any possible reads originating from the species of interest as well as related reference species (**Table 55**). Bowtie2 and the SLU custom Snakemake pipeline[209] were used for this and ended up performing the query with a database of 44 fish genomes (**Table 56**). Two of these 44 genomes are from Arctic endemic species, 5 are from Arctic-Boreal species known to occur in the CAO and 5 are from Arctic-Boreal species common in adjacent seas that could be expected to expand their distributions to the CAO. The other 32 are genomes from related species The non-artic species were included in the database to reduce the likelihood of spurious mappings.

COI-marker mapping derived from metagenomic and metatranscriptomic sequencing

The raw DNA sequence data from the 401 metagenomes (**Table 54**) was filtered and trimmed to remove adaptor sequences and low-quality regions. The reads were then mapped to the fish database using Bowtie2. Reads with good alignment scores were then used to search the NCBI nucleotide database (nt) using Blastn. Reads with good Blastn alignments to a fish species (Actinopterygii or Chondrichthyes) were considered positive hits. That is, a two-step screening procedure was used to classify reads from fish to reduce the number of false positives. The details of these steps can be found in a custom snakemake pipeline used for mapping reads to the fish genomes database and the COI database[210]. It was noted that this approach was influenced by the quality and contiguity of the genome assemblies in the database, the representation of fish taxa in NCBI nt and the composition of the query database. That means that it is expected that the pipeline is, e.g., better at detecting species with high-quality reference genomes. Read counts

---

[209] SLU custom Snakemake pipeline [https://github.com/achurcher/build_genomes_database]
[210] COI database [https://github.com/NBISweden/LTS-BiodiversityAnalysisCAO]

between species are therefore not directly comparable while differences in read counts within a species should be considered as more informative.

As a parallel to the genome-wide mapping, a track of the pipeline was developed to screen the reads from the 856 DNA metagenomes and RNA metatranscriptomes (**Table 54**) for COI markers. To achieve this, first a COI database was compiled containing ca. 2.1 million reference sequences, the majority of which were obtained from the BOLD database[211]. The reference sequences underwent quality checks to remove ambiguous nucleotides and to standardize taxonomic assignments. Finally, sequences were clustered at 100% nucleotide identity. The database was supplemented with relevant sequences missing in BOLD, including those from 77 zooplankton and 30 fish reference specimens (**Table 51**, **Table 52**).

Table 55: List of 30 fish species and 8 squid species that are known to occur in the CAO or perhaps could occur there because they occur in adjacent seas.

Order	Species	Common name	Known distribution
Gadiformes	*Arctogadus glacialis*	Ice cod	in CAO, sympagic
Gadiformes	*Boreogadus saida*	Polar cod	in CAO, sympagic, very common
Rajiformes	*Amblyraja hyperborea*	Arctic skate	in CAO, pelagic
Myctophiformes	*Benthosema glaciale*	Glacier lantern fish	in CAO, pelagic, probably common
Perciformes	*Cottunculus microps*	Polar sculpin	in CAO, pelagic
Gadiformes	*Gadus chalcogrammus*	Walleye pollock	in CAO, pelagic
Gadiformes	*Gadus morhua*	Atlantic cod	in CAO, pelagic
Perciformes	*Anisarchus medius*	Stout eelblenny	in CAO, demersal
Perciformes	*Artediellus atlanticus*	Atlantic hookear sculpin	in CAO, demersal
Perciformes	*Careproctus reinhardti*	Sea tadpole	in CAO, demersal
Perciformes	*Liparis fabricii*	Gelatinous snailfish	in CAO, demersal
Perciformes	*Lycodes adolfi*	Adolf's eelpout	in CAO, demersal
Perciformes	*Lycodes frigidus*	Glacial eelpout	in CAO, demersal
Perciformes	*Lycodes polaris*	Canadian eelpout	in CAO, demersal
Perciformes	*Lycodes sagittarius*	Archer eelpout	in CAO, demersal
Perciformes	*Lycodes seminudus*	Longear eelpout	in CAO, demersal
Perciformes	*Paraliparis bathybius*	Black seasnail	in CAO, demersal
Pleuronectiformes	*Reinhardtius hippoglossoides*	Greenland halibut	in CAO, demersal
Perciformes	*Rhodichthys regina*	Threadfin seasnail	in CAO, demersal
Gadiformes	*Melanogrammus aeglefinus*	Haddock	in CAO (marginal ice zone)
Perciformes	*Sebastes mentella*	Beaked redfish	in CAO (marginal ice zone)
Pleuronectiformes	*Hippoglossoides robustus*	Flathead sole	perhaps in CAO, but not recorded there, demersal
Pleuronectiformes	*Pleuronectes quadrituberculatus*	Alaska plaice	perhaps in CAO, but not recorded there, demersal
Rajiformes	*Amblyraja radiata*	Starry ray	perhaps in CAO, but not recorded there
Aulopiformes	*Arctozenus risso*	Spotted barracudina	perhaps in CAO, but not recorded there
Rajiformes	*Bathyraja spinicauda*	Spinetail ray	perhaps in CAO, but not recorded there
Clupeiformes	*Clupea harengus*	Atlantic herring	perhaps in CAO, but not recorded there
Gadiformes	*Gadus macrocephalus*	Pacific cod	perhaps in CAO, but not recorded there
Myctophiformes	*Lampanyctus macdonaldi*	Rakery beaconlamp	perhaps in CAO, but not recorded there

---

[211] BOLD database [boldsystems.org/]

Osmeriformes	*Mallotus villosus*	Capelin	perhaps in CAO, but not recorded there
Rajiformes	*Rajella fyllae*	Round ray	perhaps in CAO, but not recorded there
Squaliformes	*Somniosus microcephalus*	Greenland shark	perhaps in CAO, but not recorded there
Oegopsida	*Gonatus fabricii*	Armhook squid	in CAO, pelagic
Oegopsida	*Gonatus steenstrupi*	Atlantic gonate squid	perhaps in CAO, but not recorded there, pelagic
Octopoda	*Bathypolypus arcticus*	Spoonarm octopus	perhaps in CAO, but not recorded there, demersal
Octopoda	*Bathypolypus bairdii*	?	perhaps in CAO, but not recorded there, demersal
Octopoda	*Bathypolypus pugniger*	?	perhaps in CAO, but not recorded there, demersal
Octopoda	*Cirroteuthis muelleri*	?	perhaps in CAO, but not recorded there, demersal
Sepiida	*Rossia glaucopis*	Bobtail squid. (genus)	perhaps in CAO, but not recorded there, demersal
Sepiida	*Rossia moelleri*	Bobtail squid (genus)	perhaps in CAO, but not recorded there, demersal
Sepiida	*Rossia palpebrosa*	Warty bobtail squid	perhaps in CAO, but not recorded there, demersal

Table 56: List of the 44 target species in the database for direct genome mapping. These are species with Arctic-Boreal distributions and species closely related to them. * = species known to occur in the CAO but with unknown distributions. ** = additional species caught by the EFICA Consortium in the inflow of Atlantic water into the CAO (see WP5).

Fish species	Family / Order	Arctic or related	GenBank Assembly Accession
Boreogadus saida (polar cod) *	Gadidae	Arctic endemic	GCA_900302515.1
Arctogadus glacialis (ice cod) *	Gadidae	Arctic endemic	GCA_900303235.1
Benthosema glaciale (glacier lanternfish) *	Myctophidae	Arctic-Boreal	GCA_900323375.1
Gadus morhua (Atlantic cod) *	Gadidae	Arctic-Boreal	GCA_902167405.1
Gadus chalcogrammus (walleye pollock) *	Gadidae	Arctic-Boreal	GCA_900302575.1
Reinhardtius hippoglossoides (Greenland halibut) *	Pleuronectidae	Arctic-Boreal	GCA_006182925.3
Amblyraja radiata (thorny skate) *	Rajidae	Arctic-Boreal	GCA_010909765.2
Melanogrammus aeglefinus (haddock) **	Gadidae	Arctic-Boreal	GCA_900291075.1
Sebastes mentella (beaked redfish) **	Sebastidae	Arctic-Boreal	GCA_916701205.1
Gadus macrocephalus (Pacific cod)	Gadidae	Arctic-Boreal	GCA_025728055.1
Mallotus villosus (capelin)	Salangidae	Arctic-Boreal	GCA_903064625.1
Clupea harengus (Atlantic herring)	Clupeidae	Arctic-Boreal	GCA_900700415.2
Eleginus gracilis	Gadidae	related	GCA_025629765.1
Gadiculus argenteus	Gadidae	related	GCA_900302595.1
Merlangius merlangus	Gadidae	related	GCA_900323355.1
Pollachius virens	Gadidae	related	GCA_900312635.1
Trisopterus minutus	Gadidae	related	GCA_900302415.1
Hippoglossus stenolepis	Pleuronectidae	related	GCA_022539355.2
Hippoglossus hippoglossus	Pleuronectidae	related	GCA_009819705.1
Platichthys stellatus	Pleuronectidae	related	GCA_016801935.1
Verasper variegatus	Pleuronectidae	related	GCA_013332515.1
Leucoraja erinacea	Rajidae	related	GCA_000238235.1
Sebastes maliger	Sebastidae	related	GCA_916701265.1
Hozukius guyotensis	Sebastidae	related	GCA_916700995.1
Adelosebastes latens	Sebastidae	related	GCA_916700685.1
Sebastolobus alascanus	Sebastidae	related	GCA_916701645.2
Sebastiscus tertius	Sebastidae	related	GCA_916701665.1
Helicolenus hilgendorfi	Sebastidae	related	GCA_916700815.1
Alosa alosa	Clupeidae	related	GCA_017589495.2

Alosa sapidissima	Clupeidae	related	GCA_018492685.1
Limnothrissa miodon	Clupeidae	related	GCA_017657215.1
Sardina pilchardus	Clupeidae	related	GCA_003604335.1
Tenualosa ilisha	Clupeidae	related	GCA_015244755.2
Parasudis fraserbrunneri	Aulopiformes	related	GCA_900302295.1
Clinocottus analis	Cottidae	related	GCA_023055335.1
Cottus gobio	Cottidae	related	GCA_023566465.1
Cottus rhenanus	Cottidae	related	GCA_001455555.1
Myoxocephalus scorpius	Cottidae	related	GCA_900312955.1
Taurulus bubalis	Cottidae	related	GCA_910589615.1
Liparis tanakae	Liparidae	related	GCA_006348945.1
Pseudoliparis sp.	Liparidae	related	GCA_004335475.1
Protosalanx chinensis	Salangidae	related	GCA_010882115.1
Cebidichthys violaceus	Stichaeidae	related	GCA_023349555.1
Ophthalmolycus amberensis	Zoarcidae	related	GCA_024529925.1

The starting point for the COI marker mapping were sequencing reads filtered and trimmed as In the genome mapping track. These were first mapped against the COI references sequences using minimap2 and only reads aligned with high similarity (> 95%) and sufficient length (>80 bp) to at least one reference sequence were kept. These reads were then classified taxonomically using the Sintax[212] method implemented in vsearch[213] and using the COI database described above. Each read was given the most resolved assignment possible at a minimum bootstrap support of 80%. After detailed inspection of taxonomic assignments from a subset of samples it was found that the BOLD database contained COI sequences with very high similarity (up to 100%) to bacterial sequences. This led to the inclusion of "contaminant" bacterial reads in the first step of the pipeline, and the mis-assignment of these by Sintax in the second step. Therefore, an intermediate filtering step using KrakenUniq was added that removed reads with a high chance of having a prokaryotic origin. This strategy allowed us to confidently identify and assign taxonomy to COI marker reads in the metagenomes and -transcriptomes. Taxonomic assignments were then counted in each sample and all counts were collated into a combined taxonomic profile table.

12S-marker mapping derived from metagenomic and metatranscriptomic sequencing

To complement the COI marker metabarcoding-based species identification and more specifically select and describe fish and mammalian taxa, the reads from the 856 DNA metagenomes and RNA metatranscriptomes (**Table 54**) were screened for the 12S rRNA markers. A bioinformatics pipeline was developed for creating 12S reference databases from different sequence repositories. This included:

- Downloading reference sequences from different online repositories,
- Supplementing the database with sequences from the 77 zooplankton and 30 fish reference specimens (**Table 51**, **Table 52**).
- Annotating the sequences with full taxonomic metadata that will enable species identification,
- Retrieval of amplicon regions (regions that are potential targets for the PCR products) through:
  - an *in silico* PCR reaction to filter out sequences/parts of sequences that are useless in the downstream analyses as they are unlikely to match lab PCR products from the samples,
  - an additional pairwise global alignment step to retrieve amplicon sequences with missing primer-binding regions in the reference sequence
- Removal of redundant sequences from the reference database (dereplication) to improve computational efficiency of the workflow and to prevent ambiguous identifications
- Additional database cleaning and formatting
- Creation of a summary report describing reference databases (number of sequences, how many species of interest are covered, etc.)

The workflow was implemented as bash script and can be run as a standard command line program. Apart from custom-developed steps, a custom-modified version of CRABS reference database creator was used, which is available as source code[214] at To ensure full portability of the developed pipeline between different high-performance-computing hardware architectures and operating systems, a so-called Docker container was created that enables the end user to easily and painlessly use the here newly developed bioinformatic tools and workflow. The container is available at Dockerhub public

---

212 Edgar RC: SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences [https://doi.org/10.1101/074161]
213 Rognes T, et al. (2016) VSEARCH: a versatile open source tool for metagenomics. PeerJ [https://doi.org/10.7717/peerj.2584]
214 https://github.com/mkierczak/reference_database_creator

containers repository (docker pull quiestrho/crabs:1.0.0). Unfortunately, while the COI-marker mapping method succeeded, the 12S-marker mapping method did not generate any results.

COI- and 12S-amplicon mapping from DNA extraction followed by PCR amplification and MiSeq sequencing

The reads obtained from 535 DNA samples by MiSeq sequencing were processed in a usearch/vsearch-based pipeline. The script can be found at github.com/NBISweden/LTS-BiodiversityAnalysisCAO/scripts/. Reads were first merged with bbmerge, and trimmed to remove primers and discard low-quality reads with fastp. These reads were then filtered (max estimated errors 2) and dereplicated, with vsearch. The dereplicated reads are then clustered into amplicon sequence variants (ASVs) using usearch's unoise3-algorithm. Finally, the ASVs are annotated with vsearch's implementation of Sintax, using the above-mentioned COI and 12s databases. The data is then split into ASVs annotated by either of the two databases.

7.5. Results and discussion of WP7

Human contamination of the samples

The dominant species with eDNA present in the samples was *Homo sapiens*; its eDNA occurred in more than 70% of the samples and in all types of samples (water, ice habitat, sediment). This is not surprising because the molecular methods used are extremely sensitive. Contamination of the samples has probably happened mainly during sampling in the field and sample handling until they were collected on the encapsuled sterile Sterivex™ filters. Contamination in the land-based molecular laboratories where the samples were extracted is unlikely because of the strictly employed aseptic routines and inclusion of negative controls. Human contamination is common in metagenomics and amplicon studies, but since the SC07 project focuses on other organisms than humans, this issue could be handled easily by removing reads that map to the human genome.

The contamination of the Arctic environment through the release of grey water from the icebreakers was another likely source of contamination, and one which complicates the current study. DNA sequences from human as well as animals consumed on board (pig, cow and chicken) were identified in the 12S amplicon data (**Figure 88**), and even pets staying behind in the home of the cruise participants, illustrated by extremely rare but still detectable dog and cat sequences. As this study in the first place focused on fish, there is a real danger that some of the fish data in the sequence databases are "human" contaminations through fish that was consumed on-board during the expeditions. A strong recommendation for future fish eDNA studies in the CAO is thus to not serve seafood on-board.

An alternative source of "human" contamination of the Arctic environment was caused by the fishing activities by the EFICA Consortium (**Figure 88**). During the MOSAiC the bait consisted of the squid *Ilex* sp. and the shrimp *Pandalus borealis* purchased in Tromsø (northern Norway). During the SAS-Oden expedition the bait consisted of Atlantic herring *Clupea harengus* caught in the Baltic Sea, the veined squid *Loligo forbesi* and the shrimp *Pandalus borealis* purchased in Lysekil, Sweden. DNA from the Atlantic herring was detected in the genome-based and 12S amplicon approach and was abundant in samples from SAS-Oden. It was less abundant in the COI marker data. The herring signal in the data set is most likely because Atlantic herring was used as bait when fishing from the ice with lines and traps next to the ship. *Pandalus borealis* was only very rarely detected in the COI marker data sets. A strong recommendation for future fish eDNA studies in the CAO is thus to use other organisms than marine ones (e.g., from freshwater) as bait.
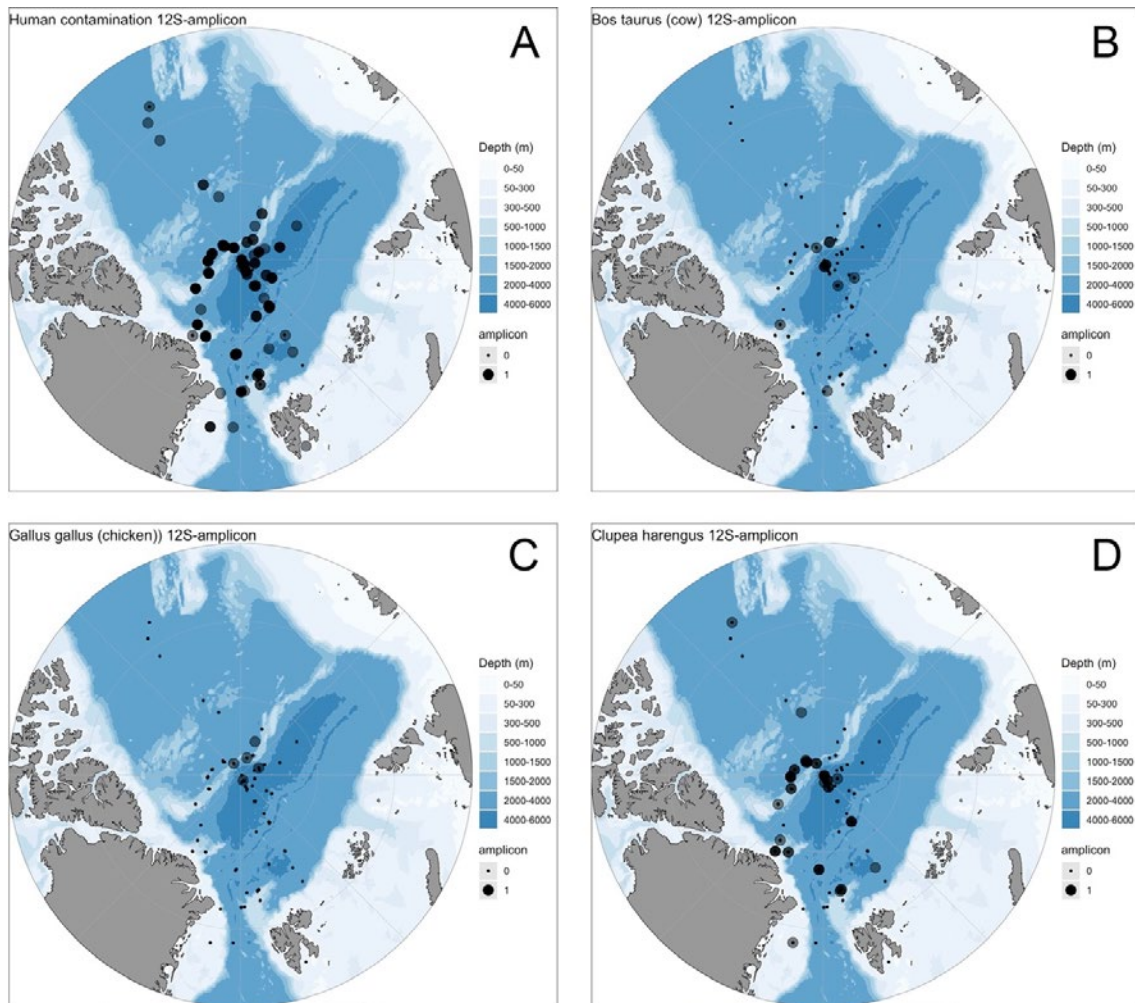
Figure 88: Examples of "human" contamination of eDNA data in the 12S amplicon data from five expeditions presented as presence / absence data. Small black dots show the positions of the sampling stations, transparent grey circles indicate detection of the 12S amplicon, and a darker circle (several transparent grey circles on top of each other) indicates that more 12S amplicons were detected at the same or a nearby station. (A) Human 12S was detected in nearly all samples. (B) Cow 12S suggests that cow was served on board. that cow. (C) Chicken 12S suggests that chicken was served on board. (D) Atlantic herring 12S shows the use of this species as bait for fishing from the ice during the SAS-Oden expedition and/or served on board, and probably it was served on board during the Beringia expedition in the Canada Basin.

Abundance of species

The results from the fish genome and COI mapping using the metagenomes and metatranscriptomes are in this report presented as "quantitative" results (average abundances per 109 reads) while the COI and 12S mapping using PCR amplification are presented as presence/absence data (i.e., present or absent in a sample). For the quantitative results it should be realised that comparisons within species can be made, but that similar comparisons between species are for several reasons associated with high uncertainty. Moreover, some of the species that should be present based on previous reports were not even detected. One reason for this could be the considerable in the degree of eDNA released to the environment from different species due to size, age, degree of shedding eDNA, etc. For example, the macrozooplankton amphipods *Themisto*

abyssorum and *Themisto libellula* should have been abundant in the zooplankton, but were surprisingly rather scarce. This could be because *Themisto* spp. (as well as other amphipods) feature a rigid exoskeletons encapsuling their bodies that likely constrain and limit the shedding of cells from their bodies. Another reason for an apparent absence could be that certain species are not represented in the databases, or that they are in fact absent. Therefore, not too much trust should be put in apparent reported absences and focus on the species that were in fact detected. Degradation of DNA due to environmental factors, such as temperature, pH, and salinity, is also highly relevant to consider in eDNA studies. In the CAO with temperatures always below 0 °C (except for the Atlantic Water Layer with temperature ca. 0-2 °C, eDNA is probably conserved longer than in warmer habitats, and this is an important aspect that needs to be investigated in future studies of eDNA in the CAO.

Results from direct genome mapping of fish

The genome mapping approach identified 19 fish (18 to the species level, 1 to the family level) in 481 metagenomes (401 from the MOSAiC and SAS-Oden expeditions (**Table 57**) and 80 from the previous LomRogIII expedition (unpublished data Pauline Snoeijs-Leijonmalm). Sixteen of these fish species were present in more than 5% of the samples while the others occurred more rarely. The latter group includes two fish species that are known to occur in the CAO (*Reinhardtius hippoglossoides* and *Amblyraja radiata*). A few reads mapped to the genomes of an additional 13 species in some samples however this is considered as background noise and is not evidence for the presence of these species in the CAO. This group of species is therefore not considered any further in the discussion.

The distribution map for the genome of Atlantic herring (**Figure 89 A**) shows that most of the identified genomes were from the ice habitat during the SAS-Oden expedition. This reflects the use of this species as bait when fishing from the ice, and probably also in the discharge grey water as herring was served for breakfast throughout the expedition. The negligible abundance of the herring genome during the MOSAiC expedition (**Table 57**) supports this conclusion. The red circles in **Figure 89 A** indicate that the herring genome was found in the pelagic zone, but analysis does not prove that Atlantic herring lives in the CAO.

The distribution of the genomes of the two Arctic gadoids *Boreogadus saida* (**Figure 89 B**) and *Arctogadus glacialis* (**Figure 89 C**) show similar patterns, although *Boreogadus* is much more abundant than *Arctogadus*. The quantitative distribution of the genomes of these two species can be compared because they are closely related and have most probably the same way of shedding eDNA to the environment. These results would confirm the hypothesis that *Boreogadus* and *Arctogadus* use the ice habitat for feeding during the transpolar drift and move to deeper water when they approach Greenland and Svalbard. In the central CAO *Arctogadus* eDNA seems to occur more in the water column than that of *Boreogadus*.

In contrast to the Atlantic herring contaminations and the Arctic endemic gadoids, the eDNA from one or several eelpouts (Zoarcidae) species basically occurred only in the water column (**Figure 89 D**). Eelpout eDNA occurred in 29 samples out of 127 collected from the water column during the SAS-Oden expedition. Since eelpouts are demersal species their DNA should not be in the ice, so the indicated habitat is likely correct. The genome of an Antarctic eelpout species *Ophthalmolycus amberensis* was identified. However, this identification is most probably wrong. It is more likely that the sequence reads are from one or several other eelpouts belonging to the family Zoarcidae. Currently, there is only one genome available from this family. Therefore, sequence reads from other eelpout species will map to the *O. amberensis* genome assembly in the absence of a true target in the database. The filtering criteria used for read assignment

was quite stringent in order to reduce cross-species read alignments. It is therefore possible that there are many more sequence reads from eelpout species in the samples but that they were not retrieved using the *O. amberensis* genome assembly and the counts from Zoarcidae are an underestimate. It can be concluded that further genome sequencing of eelpout species is warranted and that the current metagenomic data set can be revisited and re-analysed when such resources are available.

Table 57: Results from direct genome mapping, showing 19 fish genomes identified in the 481 metagenomes (DNA) and their occurrence in the whole data set (% of samples). Another, 13 species with lower abundances were deleted and can be considered background noise. The expedition data shown are average abundances per 10^9 reads for each expedition/habitat combination. All expeditions took place in the CAO, but MOSAiC was also partly in the inflow area of Atlantic water to the CAO in Fram Strait and SAS-Oden was also partly in the outflow area of Atlantic water from the on the Greenland shelf. The habitats are: ice (snow, melted ice, brine, melt ponds, and water at the ice-seawater interface), water (seawater between 11 and 4500 m of depth), and deep-sea sediment (only sampled during the SAS-Oden expedition).

Average library size per sample:		399 397 704	366 488 294	240 208 942	322 824 018	253 149 915
Number of samples:		46	56	156	16	127
Fish species	% of samples	MOSAiC ice	MOSAiC water	SAS-Oden ice	SAS-Oden water	SAS-Oden sediment
Clupea harengus	62.8	0.1	0.3	7837.6	962.1	14.7
Boreogadus saida	51.1	558.2	599.3	58.2	87.3	5.4
Melanogrammus aeglefinus	36.8	163.7	22.2	11.4	24.6	
Gadus morhua	30.1	42.9	30.0	9.4	7.7	
Sebastes mentella	21.2	3.0	0.8	5.6	19.3	
Arctogadus glacialis	11.6	7.0	7.5	0.4	2.4	
Merlangius merlangus	11.4	0.8	1.0	5.7	3.1	
Pollachius virens	8.9	9.6	7.5	0.2	0.6	
Ophthalmolycus amberensis	7.3	0.1	0.6	0.0	31.1	
Gadus macrocephalus	6.9	4.7	3.3	0.2	1.4	
Gadus chalcogrammus	6.4	2.7	2.1	0.1	0.7	
Platichthys stellatus	5.0	0.1		0.5	0.8	
Eleginus gracilis	3.1	1.3	0.7	0.1	0.3	
Mallotus villosus	3.1	0.1	0.3	0.2	1.5	
Pseudoliparis sp.	2.3	0.1	0.1		1.2	
Trisopterus minutus	2.3	0.7	0.8	0.1		
Amblyraja radiata	1.7				1.5	
Myoxocephalus scorpius	1.7			0.3	0.1	
Reinhardtius hippoglossoides	1.5		0.2	0.04	0.2	

Figure 89: Distribution maps of fish genomes expressed as number of genomes per 10⁹ reads. Small black dots show the positions of the sampling stations, yellow circles show the abundance of the fish genome in sea ice habitats, red circles show the abundance of the fish genomes in the water column. (A) Atlantic herring *Clupea harengus*. (B) Polar cod *Boreogadus saida*. (C) Ice cod *Arctogadus glacialis*. (D) Eelpout (Zoarcidae).

One of the major questions of the SC07 project is how far Atlantic species could protrude into the CAO? The eDNA results suggest that several Atlantic species could have a broad distribution in CAO and the Greenland shelf than previously assumed, even if they were not caught during the SAS-Oden expedition. In the inflow of Atlantic water to the CAO, in fact three Atlantic species were caught: Atlantic cod *Gadus morhua*, haddock *Melanogrammus aeglefinus* and beaked redfish *Sebastes mentella* (see WP5). In this study the DNA of these three species was identified far into the CAO (**Figure 90 A-C**). An addition to this list of protruding Atlantic species is whiting *Merlangius merlangus* (**Figure 90 D**). Higher abundances of Atlantic cod eDNA in the water column were found exactly in the two areas where this species was caught, near the Yermak Plateau in the Atlantic inflow and on the eastern side of the Amundsen Basin (see WP5). A third area with higher abundances of eDNA from this species was at the Lomonosov Ridge close to the Greenland shelf and this could suggest that Atlantic cod could live there. A rather surprising finding of high abundance of eDNA from Atlantic Cod was in the ice on the western side of the Amundsen Basin. Without any direct evidence, it can be suggested that this his could have been an occasion when Atlantic cod was served onboard the ship with eDNA contamination from the grey water. The other three Atlantic species showed

similar patterns but without this patchy eDNA occurrence in the eastern Amundsen Basin, and were even more distributed across the western Lomonosov Ridge than Atlantic cod.
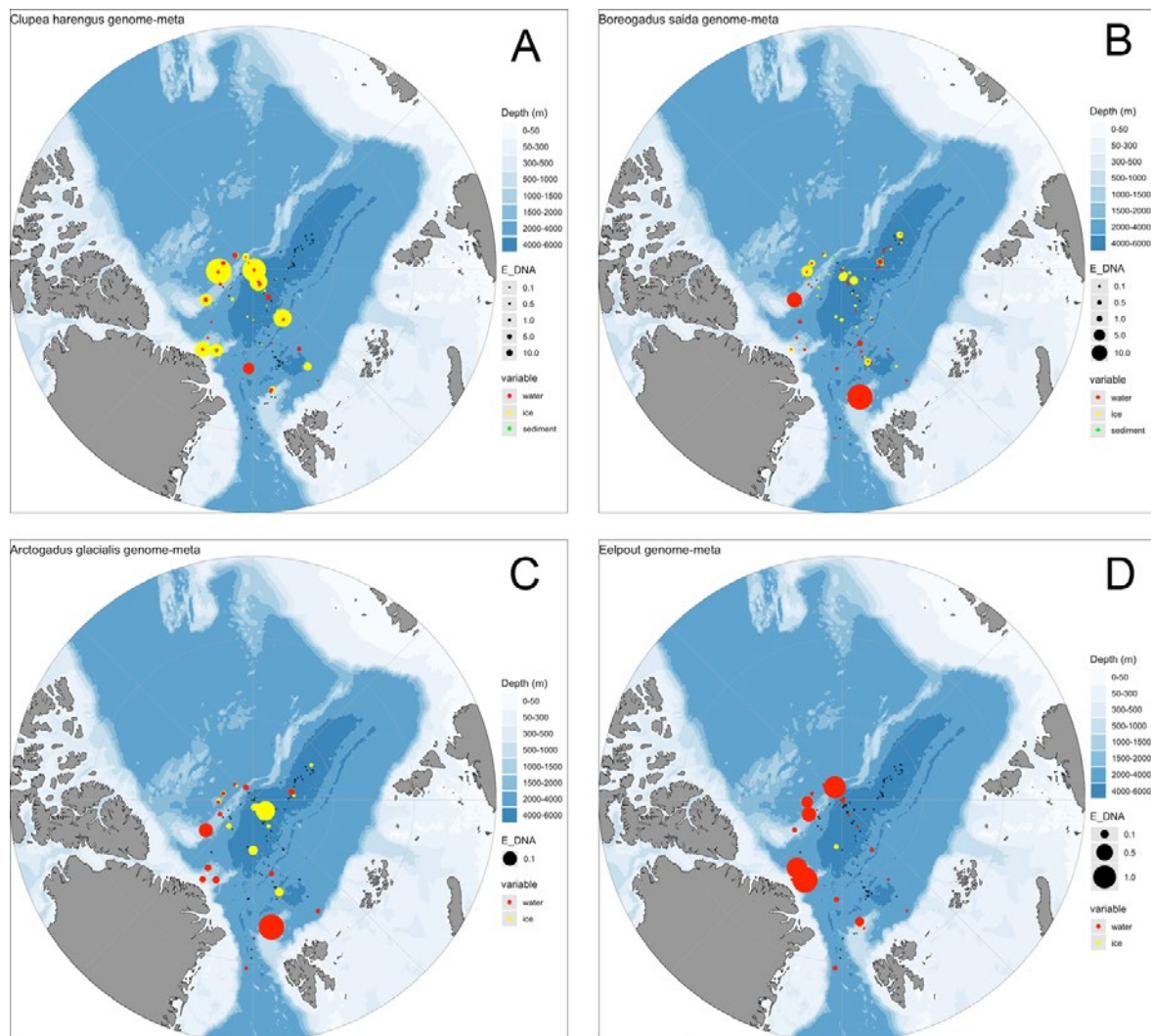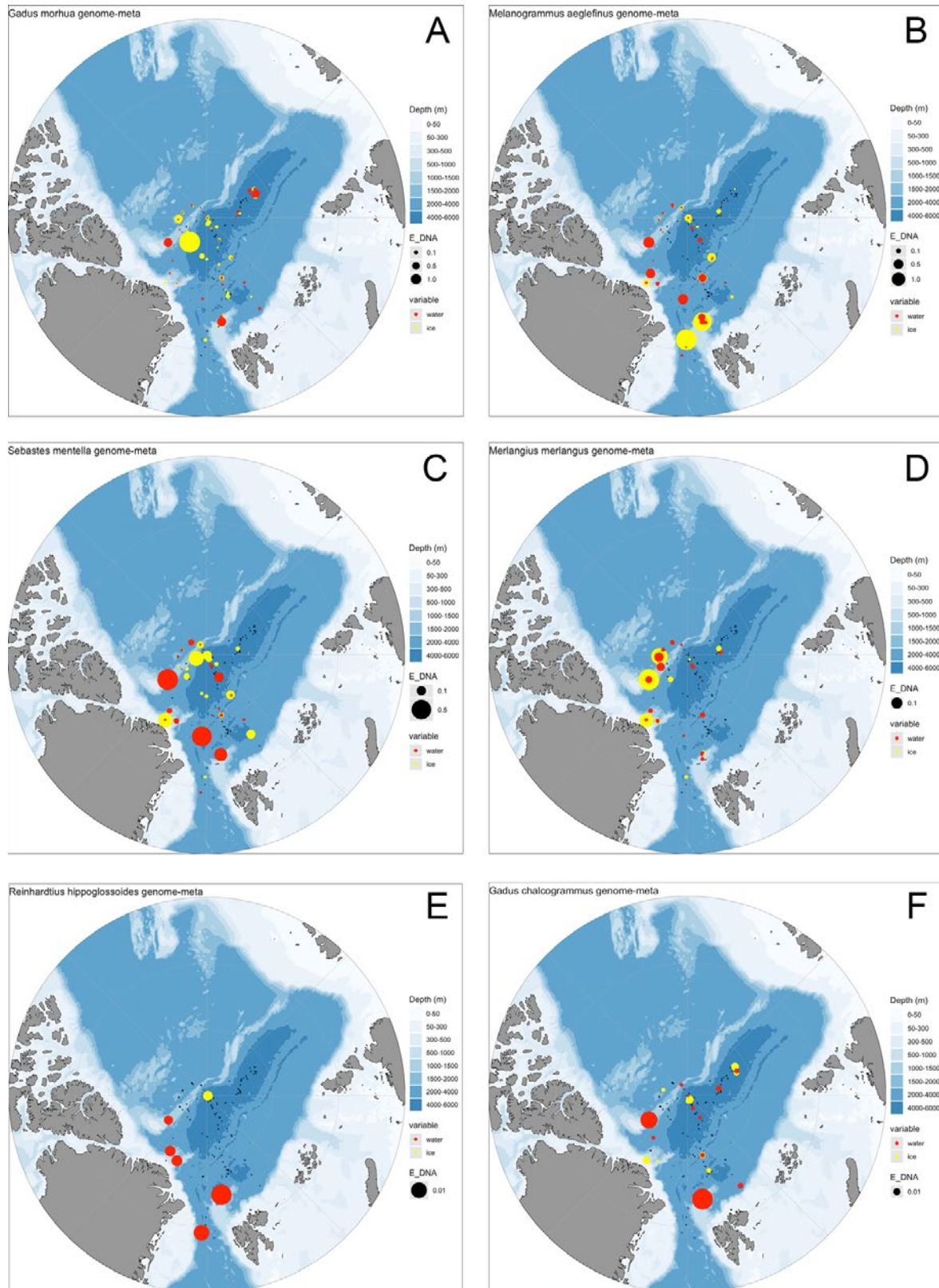

Figure 90: Distribution maps of fish genomes expressed as number of genomes per 109 reads. Small black dots show the positions of the sampling stations, yellow circles show the abundance of the fish genome in sea ice habitats, red circles show the abundance of

the fish genomes in the water column. (A) Atlantic cod *Gadus morhua*. (B) Haddock *Melanogrammus aeglefinus*. (C) Beaked redfish *Sebastes mentella*. (D) Whiting *Merlangius merlangus*. (E) Greenland halibut *Reinhardtius hippoglossoides*. (F) Walleye pollock *Gadus chalcogrammus*.

The flatfish Greenland halibut which is a demersal species, featured an eDNA distribution that was restricted to more shallow waters (**Figure 90 E**). For this species, there is a disturbing spot in the ice habitat at the North Pole and Greenland halibut was most probably not served for dinner onboard. The species level resolution for right-eyed flounders is difficult because few sequence reads have been retrieved. The genome assemblies for righteye flounders appear to be relatively complete and there are lots of sequences from this family in GenBank so this is not likely a factor in this case.

The Pacific species walleye pollock *Gadus chalcogrammus* (**Figure 90 F**) featured an eDNA distribution pattern that was most similar to the two Arctic endemic gadoids *Boreogadus saida* and *Arctogadus glacialis*. While this is an exciting finding, it is possible that there is confusion in the identifications of *Gadus chalcogrammus* and *Arctogadus glacialis* in the taxonomic databases. This is a task for taxonomists to solve by studying the type specimens of these two species.

Unfortunately, the DNA of *Benthosema glaciale* was not identified in the data set based on the genome mapping. This species (or a close relative) was observed on the FishCam during the MOSAiC expedition (see WP3). This is a species where the genome assembly is quite fragmented and there are very few sequences from related species. However, in the second filtering step of the pipeline (the blast step) there was a read count of 2 for one sample (SAS-Oden2021-DNA-574), and while this could be a true positive, the limited data calls for caution in making any far reaching conclusions. For other samples, the mapped read counts relative to the number of locations in the genome is still close to what could be contamination in the assembly. So, at the moment it can be concluded that at this stage there is no evidence for widespread distribution of *Benthosema glaciale* in the metagenome data.

In general, the bioinformatics approach employed for the genome-based analysis appears to be highly species-specific and robust. The approach, including the database build, is calibrated to reduce the likelihood of false positives, i.e., to reduce the possibility of incorrect read assignment. While the possibility that some reads within the results have been incorrectly assigned cannot be excluded, it is much more likely that the approach used has resulted in an underestimate of true target species abundances. Moreover, it should be stressed that this approach focused on correctly identifying a subset of key fish species found in the Arctic. It is therefore not a survey of all possible Arctic fish species. Furthermore, for species with less contiguous genome assemblies (e.g., *Boreogadus saida* and *Benthosema glaciale*) and for species where there are few sequences from related species in GenBank/NCBI nr (e.g., *Benthosema glaciale*), it is more difficult to find evidence of presence based on mapping at genome-wide level.

Results from COI marker mapping

For the COI marker mapping 481 metagenomes were used, 401 from the MOSAiC and SAS-Oden expeditions (**Table 54**) and 80 from the previous LomRogIII expedition (unpublished data Pauline Snoeijs-Leijonmalm), as well as 455 metatranscriptomes from the MOSAiC and SAS-Oden expeditions (**Table 54**). Altogether, 936 metagenomes/metatranscriptomes were screened and most of the COI reads (89%) belonged to the Chromista while 7% were unidentified animals (**Figure 91**). Only 3.5% of the COI sequences were associated with animals identified at genus or species level, 0.1% fish and 3.4% invertebrates. The 31 most abundant animal species identified by their COI markers were shortlisted (**Table 58**). The eDNA of these species occurred in

812 of the 936 samples and consisted of seven nekton species (five fish and two cephalopods) and 24 zooplankton species.
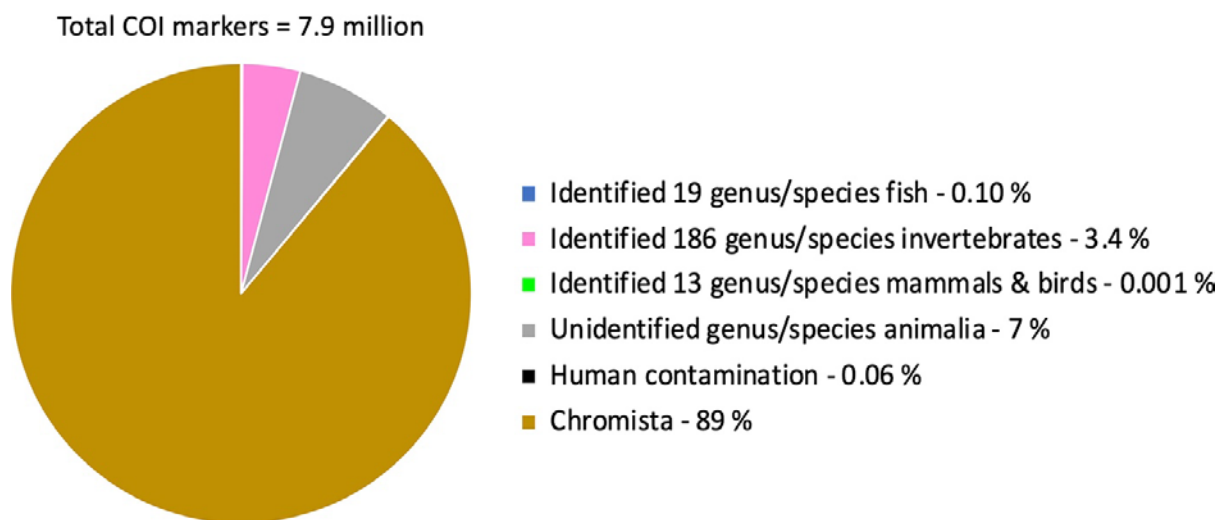
Figure 91: The composition of the COI markers detected in 936 metagenomes/ metatranscriptomes. The largest portion of COI markers detected (89%) belonged to the Chromista (photosynthetic eukaryotes). The two other major groups were unidentified animals (7%) and identified invertebrates (3.4%). In the latter group the zooplankton was represented. Fish and mammals & birds COI markers made only up less than 0.1 and 0.001%, respectively, of all COI markers detected.

The COI marker method identified – as expected – less fish records than the genome method. Altogether 19 species were detected and the taxonomic composition overlapped between the two methods. The most abundant fish species identified by this method were polar cod *Boreogadus saida*, walleye pollock *Gadus chalcogrammus*, Atlantic cod *Gadus morhua*, haddock *Melanogrammus aeglefinus*, and Atlantic herring *Clupea harengus*. The "human" contamination with Atlantic herring was much less prominent as compared to the genome-wide mapping method (compare **Table 57** and **Table 58**).

The eDNA distribution maps for polar cod and walleye pollock (**Figure 92 A,B**) did not show enhanced pelagic abundances near Greenland and Svalbard as observed with the genome mapping (**Figure 89 B**, **Figure 90 F**). However, the eDNA distribution patterns of Atlantic cod (**Figure 92 C**) and haddock (**Figure 92 D**) were similar to the actual fish catches (see WP5), but again not showing higher abundance near Greenland or Svalbard.

With this method it was expected to identify the DNA of the cephalopod mollusc armhook squid *Gonatus fabricii*, a species that was recorded with the MOSAiC FishCam (see WP3) and identified in the movies by internationally renowned cephalopod experts. Instead, the COI database identified the closely related Atlantic gonate squid *Gonatus steenstrupi*. Again, there is possibly uncertainties concerning these two species in the databases because they are very alike, and further genomic data and taxonomic studies and possibly database curation is necessary. The shortfin squid *Illex argentines* which is most prevalent along the coasts of South America, in Brazil and Argentina, is probably a misidentification by the database as well, and could represent another species of the genus *Ilex*.

The eDNA distribution pattern of *Gonatus* in the pelagic zone (**Figure 92 E**) perfectly matches the observations of squid on the FishCam (see WP3). The eDNA distribution pattern of *Ilex* (**Figure 92 F**) is a bit similar, but more difficult to explain. Perhaps two squid species occur in the CAO or maybe the databases confused even the identification of this squid with *Gonatus*. It should, however, be considered that these squids reproduces by the gravid female being eaten from the inside by her growing offspring during which the dying female body floats to the ocean surface. This could explain the occurrence of pelagic squid eDNA in the ice habitat (**Figure 92 F**).

Table 58: Results from COI mapping, showing the 31 species with the highest abundances in the 812 metagenomes (DNA) and metatranscriptomes (RNA) in which COI reads were identified and their occurrence in the whole data set (% of samples). All expeditions took place in the CAO, but MOSAiC was also partly in the inflow area of Atlantic water to the CAO in Fram Strait and SAS-Oden was also partly in the outflow area of Atlantic water from the on the Greenland shelf. The habitats are: ice (snow, melted ice, brine, melt ponds, and water at the ice-seawater interface), water (seawater between 11 and 4500 m of depth), and deep-sea sediment (only sampled during the SAS-Oden expedition). Taxa ending in "_X" represent hits to references for which no consensus species name could be assigned in the database.

Average library size:		178866390	165980827	239331293	250816180
Number of samples:		225	305	112	123
Species	Occurrence (% samples)	MOSAiC ice	MOSAiC water	SAS-Oden ice	SAS-Oden water
CHORDATA - FISH					
Boreogadus saida	35.8	211.6	26.4		
Gadus morhua	5.7	2.9	1.4		
Gadus chalcogrammus	2.7	1.1	0.2	0.1	
Clupea harengus	2.1		0.01	2.7	0.1
Melanogrammus aeglefinus	1.4	0.5	0.2		
MOLLUSC - CEPHALOPODA					
Illex argentinus	24		0.9	0.5	0.1
Gonatus steenstrupi	10		0.1	0.4	
CHORDATA - APPENDICULARIA					
Fritillaria crassifolia	17.5	4.2	2.9	25.5	2.2
ARTHROPODA					
Oithona similis	55.2	1972.4	1453.0	25.9	31.7
Microsetella norvegica	27.5	896.9	12.8	31.9	1.2
Calanus hyperboreus	19.6	166.7	88.0	0.4	18.2
Calanus glacialis	18.1	37.0	459.3	1.5	1.3
Triconia borealis	15.6	26.0	178.2	0.1	5.4
Metridia longa	9.0	0.6	134.9	0.0	16.0
Microcalanus pygmaeus	8.0	14.0	368.7		6.4
Apherusa_X	7.9	5.7	0.5	2.7	0.1
Calanus finmarchicus	7.8	24.9	11.4	0.1	4.1
Apherusa glacialis	7.6	7.0	0.8	8.5	0.6
Oncaea_X	5.4	1.3	66.3		16.4
Microcalanus pusillus	3.8	14.0	16.8		0.7
Cyclocaris guilelmi	3.6	0.0	93.9		
Pseudocalanus minutus	2.6	30.1	2.5	0.4	
Paraeuchaeta glacialis	2.2	1.7	503.2		0.1
Spinocalanus longicornis	1.2		24.1		0.0
Oithona atlantica	0.6		84.1		0.9
Neomormonilla minor	0.6		27.4		0.3
CNIDARIA					

Marrus orthocanna	13.7	3.4	14.6		2.3
Botrynema brucei	8.0	2.6	20.3	0.1	3.2
CHAETOGNATHA					
Eukrohnia hamata	26.7	63.9	625.5		0.9
Pseudosagitta maxima	5.3	5.1	351.0		
NEMATODA					
Halomonhystera disjuncta	1.6	2.5	11.8		
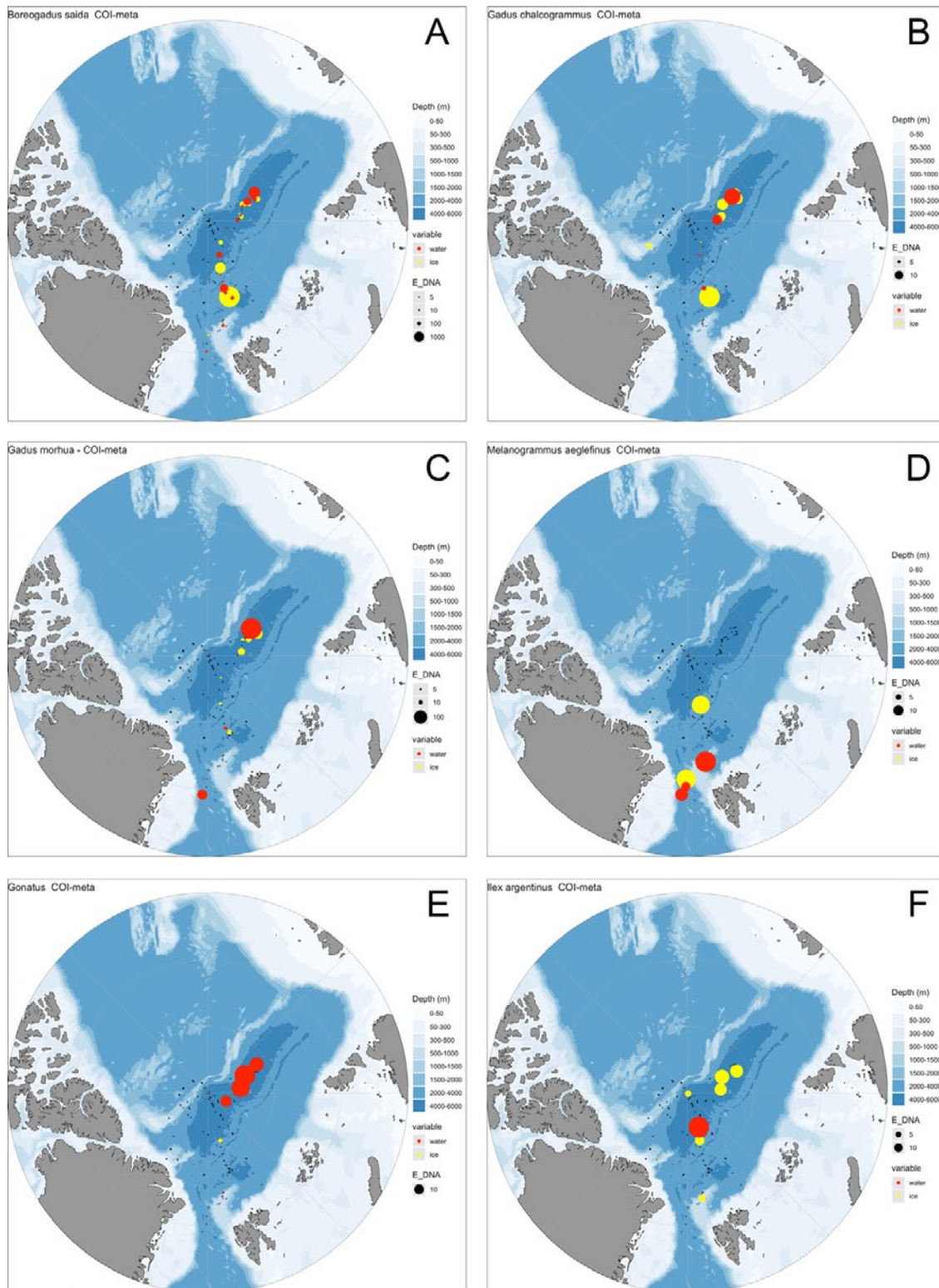
Figure 92: Distribution maps of COI markers from metagenomes and metatranscriptomes expressed as number of markers per 10^9 reads. Small black dots show the positions of the sampling stations, yellow circles show the abundance of the fish genome in sea ice habitats, red circles show the abundance of the fish or squid genomes in the water column. (A) polar cod *Boreogadus saida*. (B) Walleye pollock *Gadus chalcogrammus*. (C) Atlantic cod *Gadus morhua*. (D) Haddock *Melanogrammus aeglefinus*. (E) *Gonatus* (squid). (F) *Ilex* (squid).

The dominant zooplankton species identified with the COI marker method were all species known to be common in the CAO: the copepods *Calanus finmarchicus*, *Calanus glacialis*, *Calanus hyperboreus*, *Metridia longa*, *Microcalanus pusillus*, *Microcalanus pygmaeus*, *Microsetella norvegica*, *Oithona similis*, *Oncaea* sp., *Paraeuchaeta glacialis*, *Pseudocalanus minutus*, *Triconia borealis*, and the amphipods *Apherusa glacialis* and *Cyclocaris guilelmi*, the chaetognaths *Eukrohnia hamata* and *Pseudosagitta maxima*, the cnidaria *Marrus orthocanna* (siphonophore) and *Botrynema brucei* (hydrozoa), and the small swimming mollusc *Limacina helicina*. However, some notorious Arctic invertebrates, such as the amphipods *Eusirus holmii*, *Lanceola clausii*, *Themisto abyssorum* and *Themisto libellula*, the decapod *Hymenodora glacialis*, and the ctenophores *Beroe* sp. and *Mertensia* sp. are missing from this list (see WP4). The eDNA distributions of the zooplankton did not show any clear patterns.

Human COI reads occurred in 55% of the samples, and these samples also contained DNA traces of their typical food items (pig, cow, chicken). Marine mammals were extremely rare in the data set, but narwhal *Monodon monoceros*, walrus *Odobenus rosmarus*, harp seal *Pagophilus groenlandicus*, and ringed seal *Pusa hispida* were nevertheless identified in a few samples.

Results from 12S marker mapping

12S marker mapping used the same data set as COI marker mapping: 936 metagenomes / metatranscriptomes were screened, but no results were obtained. Then a test was made with a sample to which 12S markers from the database were added, and these were detected. From this the can be concluded that there were no 12S markers that could be identified in the shotgun sequenced metagenome/metatranscriptome data sets.

Total COI amplicons = 53.2 million

- Identified genus/species fish - 0.7 %
- Identified genus/species invertebrates - 3.4 %
- Identified genus/species mammals & birds - 0.1 %
- Unidentified genus/species animalia - 33 %
- Human contamination - 3.7 %
- Chromista - 59 %
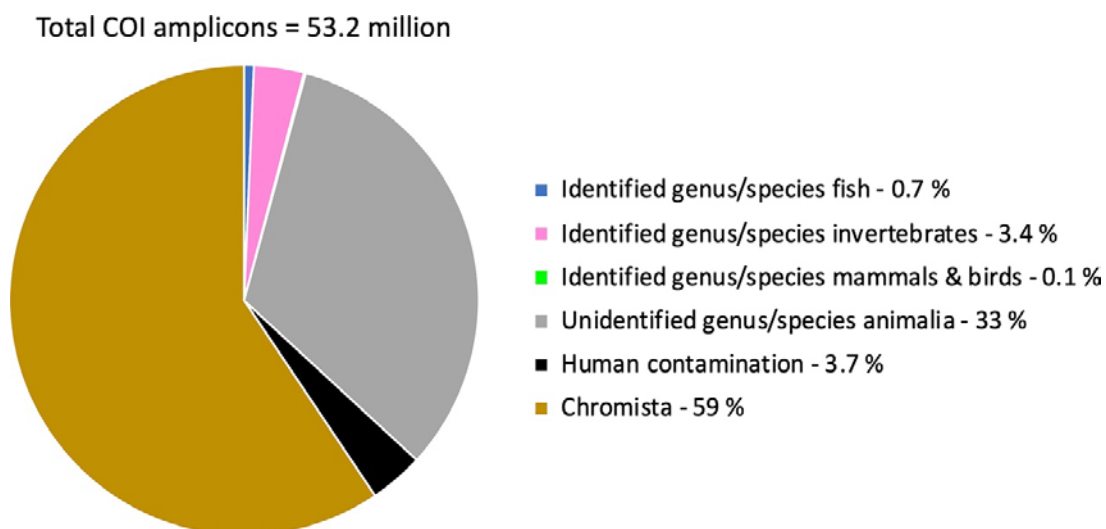
Figure 93: The composition of the COI markers detected in 652 amplicon samples. The largest portion of COI markers detected (59%) belonged to the Chromista (photosynthetic eukaryotes). The other major groups consisted of unidentified animals (33%), human contamination (3.7%), and identified invertebrates (3.4%). In the latter group the zooplankton was represented. Fish and mammals & birds COI markers made only up less than 0.7 and 0.1%, respectively, of all COI markers detected.

Results from the COI-amplicon analyses

The COI-amplicon analyses used 535 samples from the MOSAiC and SAS-Oden expeditions and (**Table 54**) and 117 samples from the previous AO02, Beringia, and LomRogIII expeditions (unpublished data Pauline Snoeijs-Leijonmalm). Altogether, 652 samples were screened analogously to the metagenome/metatranscriptome mapping. Most of the COI amplicons (59%) belonged to the Chromista and 33% were unidentified animals (**Figure 93**). Only 4.2% of the COI amplicon sequences were from animals identified to the genus or species level, 0.7% fish and 3.4% invertebrates. The 31 most abundant animal species identified by their COI markers were shortlisted (**Table 59**). eDNA of these species occurred in 431 of the 652 samples and consisted of four fish and 27 zooplankton species. The fish were polar cod *Boreogadus saida*, haddock, *Melanogrammus aeglefinus*, Atlantic cod *Gadus morhua* and a redfish identified only to

Table 59: Results from the COI amplicon analyses, showing the 31 species with the highest abundances in the 431 samples (DNA) in which COI amplicons were identified and their occurrence in the whole data set (number of samples and % of samples).

Number of samples:		430	430	251	125	54
Number of reads:		6 712 989	6 712 989	3 968 835	2 006 344	737 810
Number of COI amplicons:		4 214 100	4 214 100	2 477 300	1 414 600	322 200
Group	Genus / species	Number of samples	Occurrence (% of all samples)	Occurrence (% of ice samples)	Occurrence (% of water samples)	Occurrence (% of sediment samples)
Fish	*Boreogadus saida*	84	19.5	15.5	22.4	31.5
Fish	*Melanogrammus aeglefinus*	4	0.9	1.2	0.8	
Fish	*Gadus morhua*	1	0.2	0.4		
Fish	*Sebastes_X*	1	0.2	0.4		
Zooplankton, copepod	*Microsetella norvegica*	61	14.2	21.9	4.0	1.9
Zooplankton, copepod	*Oithona similis*	61	14.2	10.0	28.8	
Zooplankton, copepod	*Calanus hyperboreus*	28	6.5	4.4	13.6	
Zooplankton, copepod	*Cyclopoida_XXX*	17	4.0	6.4	0.8	
Zooplankton, copepod	*Calanus glacialis*	11	2.6	2.4	4.0	
Zooplankton, ostracod	*Halocyprida_XXX*	9	2.1	0.8	5.6	
Zooplankton, copepod	*Metridia longa*	8	1.9	1.2	4.0	
Zooplankton, amphipod	*Apherusa*	6	1.4	1.6	0.8	1.9
Zooplankton, hydrozoan	*Botrynema brucei*	6	1.4	2.4		
Zooplankton, chaetognath	*Eukrohnia hamata*	5	1.2	0.4	3.2	
Zooplankton, amphipod	*Cyclocaris guilelmi*	4	0.9	0.8	1.6	
Zooplankton, copepod	*Oncaea_X*	4	0.9		3.2	
Zooplankton, hydrozoan	*Marrus orthocanna*	4	0.9		3.2	
Zooplankton, copepod	*Microcalanus pusillus*	3	0.7	0.8	0.8	
Zooplankton, amphipod	*Lanceola clausii*	2	0.5	0.8		
Zooplankton, amphipod	*Onisimus nanseni*	2	0.5	0.8		
Zooplankton, amphipod	*Themisto libellula*	2	0.5	0.8		
Zooplankton, copepod	*Paraeuchaeta*	2	0.5	0.4	0.8	
Zooplankton, branchiopod	*Evadne nordmanni*	1	0.2		0.8	
Zooplankton, chaetognath	*Pseudosagitta maxima*	1	0.2	0.4		
Zooplankton, flatworm	*Fasciolopsis buski*	1	0.2	0.4		
Bird, dove	*Turtur afer*	9	2.1	0.4	6.4	
Bird, chicken	*Gallus gallus*	3	0.7	0.8		1.9
Mammal, human	*Homo sapiens*	311	72.3	71.3	63.2	98.1
Mammal, pig	*Sus scrofa*	11	2.6	4.0	0.8	
Mammal, cow	*Bos taurus*	6	1.4	1.6	0.8	1.9
Plant, red macroalga	*Neoporphyra haitanensis*	3	0.7	0.4	1.6	

genus level, *Sebastes*_X (probably beaked redfish). The zooplankton composition was similar to that of the COI marker screening of the metagenomes / metatranscriptomes (compare **Table 58** and **Table 59**). The composition was similar to previous results and did not show any clear eDNA distribution pattern. Single eDNA occurrences can represent artefacts.

The eDNA distribution maps of the four fish species showed again the dominance of polar cod *Boreogadus saida* in the CAO, now also including the Canada Basin covered by the Beringia expedition in 2005 (**Figure 94 A**). Atlantic cod *Gadus morhua* had only one record, in the outflow region of Arctic water to the North Atlantic covered by the AO02 expedition in 2002 (**Figure 94 B**). Haddock *Melanogrammus aeglefinus* had four records, three in the Atlantic water inflow area (one from MOSAiC two from AO02) and one at the Lomonosov Ridge (Beringia), while redfish *Sebastes* had only one record, also at the Lomonosov Ridge (Beringia) (**Figure 94 C,D**). The records at the Lomonosov ridge represent only one eDNA occurrence for each species and could be artefacts while the observation was also corroborated by 12S results from the same samples (**Figure 95**). No squid was discovered with this method.
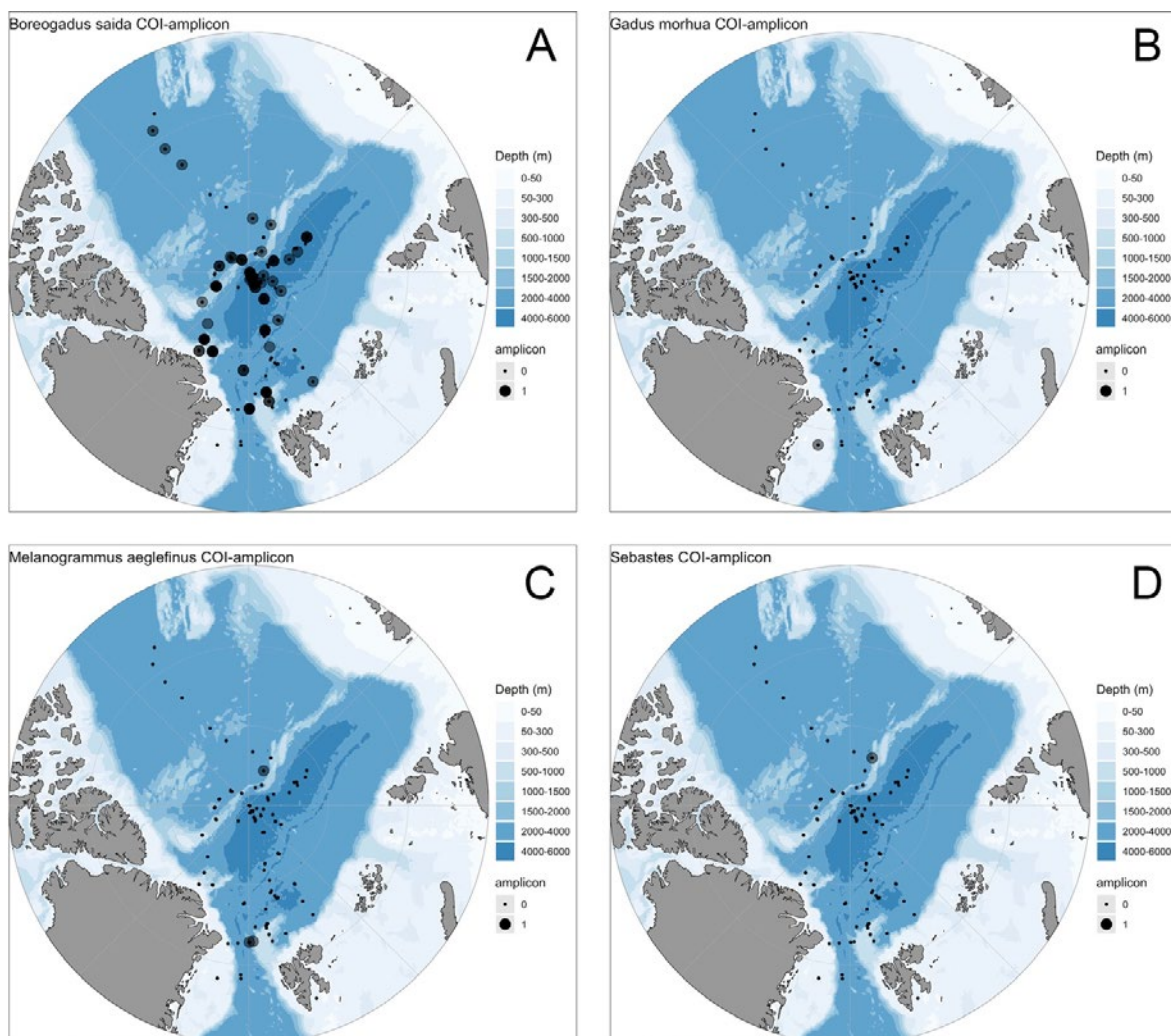

Figure 94: Distribution maps of COI amplicons presented as presence / absence data. Small black dots show the positions of the sampling stations, transparent grey circles indicate detection of the COI amplicon, and a darker circle (several transparent grey circles on top of each other) indicates that more COI amplicons were detected at the

same or a nearby station. (A) Polar cod *Boreogadus saida*. (B) Atlantic cod *Gadus morhua*. (C) Haddock *Melanogrammus aeglefinus*. (D) Redfish *Sebastes_X*.

Results from the 12S-amplicon analyses

The 12S-amplicon analyses was applied to 535 samples from the MOSAiC and SAS-Oden expeditions and (**Table 54**) and 117 samples from the previous AO02, Beringia, and LomRogIII expeditions (unpublished data Pauline Snoeijs-Leijonmalm). Altogether, 652 samples were screened and amplicons occurred in 431 of the 652 samples. The method detected the 12S amplicons of fish and mammals while invertebrates were absent.

The eDNA distribution maps of the fish species showed again the dominance of polar cod *Boreogadus saida* in the CAO, including the Canada Basin covered by the Beringia expedition in 2005 (**Figure 95 A**). Possibly the Atlantic species have wider distributions in the CAO than previously assumed. Atlantic cod *Gadus morhua* had 48 records with a wide eDNA distribution in the CAO, including three in the Canada Basin, but surprisingly not in the eastern Amundsen Basin where this species was caught in the CAO (see WP5) (**Figure 95 B**). The Atlantic cod eDNA occurred even in waters overlying the deep basins. Haddock *Melanogrammus aeglefinus* had 14 records, and *Sebastes mentella* 8 records, most of them were not over the deep basins, but at the inflow of Atlantic water to the CAO and at the Lomonosov Ridge (**Figure 95 C,D**). These findings agree with the findings of their genomes at the Lomonosov Ridge (**Figure 90**).

For the demersal fish of the CAO, the 12S method identified snailfish (Liparidae) (**Figure 95 F**), which were not detected with the genome mapping method. The genome mapping method, on the other hand, identified eelpouts (Zoarcidae) (**Figure 89 D**). Such discrepancies do not depend on the methods *per se*, but on the availability of genomes and 12S sequences of these species in the public databases.

The single record of *Gasterosteus aculeatus* (**Table 60**) probably represents a database artefact since this is a freshwater species, but the single record of *Pollachius virescens* could be real as this species was also identified by genome mapping (**Table 57**), as well as the detection of walrus *Odobenus rosmarus* in a single sample. Human contamination in the form of food and even pets was apparent in many of the samples.

Table 60: Results from the 12S amplicon analyses, showing the 17 species with the highest abundances in the 431 samples (DNA) in which 12S amplicons were identified and their occurrence in the whole data set (number of samples and % of samples).

Number of samples:		202	202	111	74	17
Number of reads:		3 380 236	3 380 236	2 043 196	1 105 828	231 212
Number of 12S amplicons:		24080	24080	20318	2335	1427
Group	Genus / species	Number of samples	Occurrence (% of all samples)	Occurrence (% of ice samples)	Occurrence (% of water samples)	Occurrence (% of sediment samples)
Fish	*Boreogadus saida*	123	60.9	46.8	75.7	88.2
Fish	*Clupea harengus*	49	24.3	34.2	14.9	
Fish	*Gadus morhua*	48	23.8	26.1	23.0	11.8
Fish	*Melanogrammus aeglefinus*	14	6.9	10.8	2.7	
Fish	*Sebastes mentella*	8	4.0	2.7	6.8	
Fish	Liparidae	4	2.0	0.9	4.1	
Fish	*Gadus chalcogrammus*	2	1.0	0.9		5.9
Fish	*Pollachius virescens*	1	0.5	0.9		
Fish	*Gasterosteus aculeatus*	1	0.5	0.9		
Bird, chicken	*Gallus gallus*	2	1.0	1.8		
Bird, turkey	*Turtur afer*	5	2.5	4.5		

Mammal, human	*Homo sapiens*	157	77.7	87.4	63.5	76.5
Mammal, pig	*Sus scrofa*	48	23.8	27.9	21.6	5.9
Mammal, cow	*Bos taurus*	13	6.4	7.2	5.4	5.9
Mammal, dog	*Canis lupus*	4	2.0	2.7		5.9
Mammal, cat	*Felis catus*	1	0.5	0.9		
Mammal, walrus	*Odobenus rosmarus*	1	0.5	0.9		
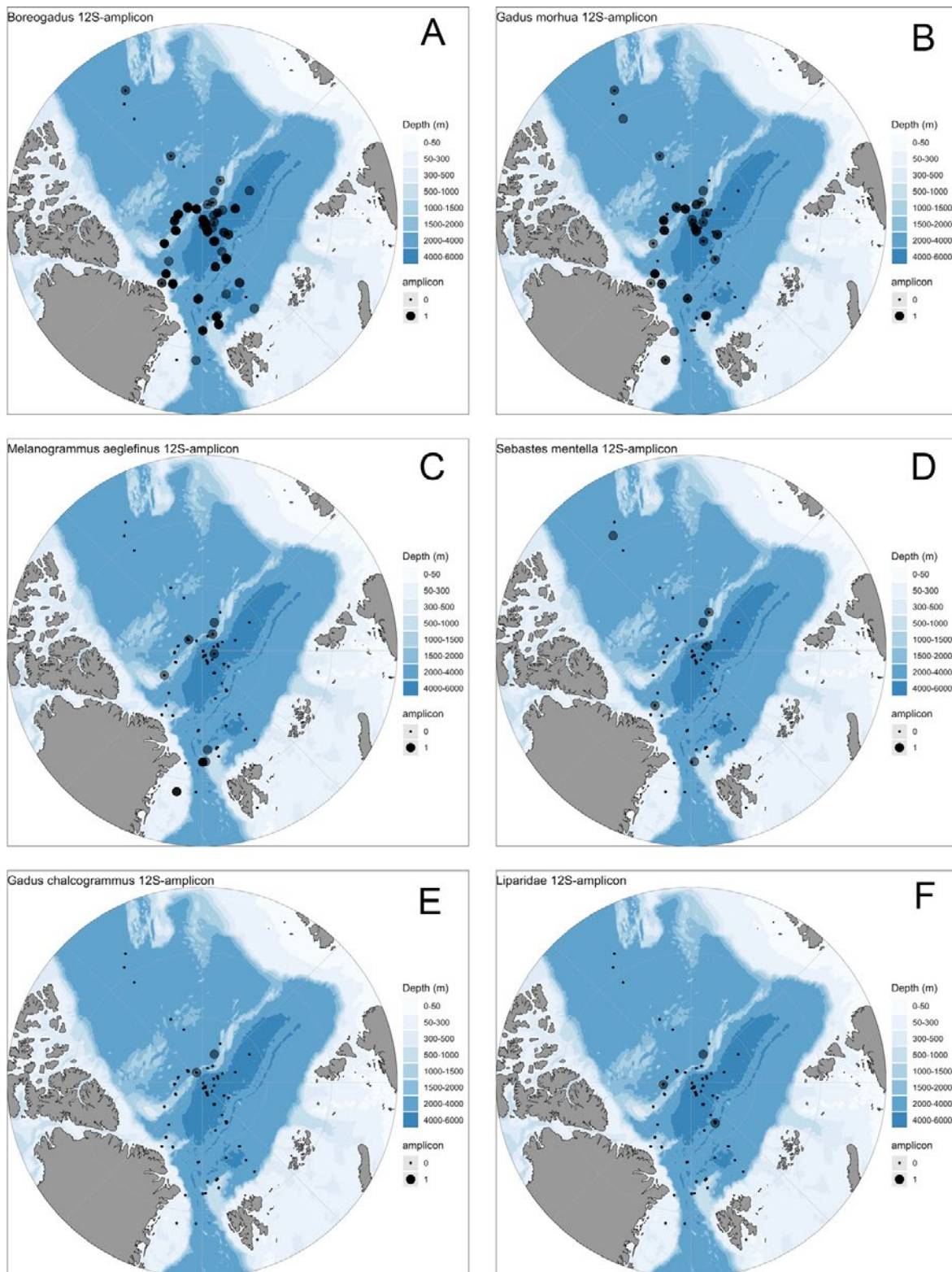
Figure 95: Distribution maps of 12S amplicons presented as presence / absence data. Small black dots show the positions of the sampling stations, transparent grey circles indicate detection of the 12S amplicon, and a darker circle (several transparent grey circles on top of each other) indicates that more 12S were amplicons detected at the same or a nearby station. (A) Polar cod *Boreogadus saida*. (B) Atlantic cod *Gadus morhua*. (C) Haddock *Melanogrammus aeglefinus*. (D) Beaked redfish *Sebastes mentella*. (E) Walleye pollock *Gadus chalcogrammus*. (F) Snailfish (Liparidae).

7.6. Answers to the WP7 research questions

(1) What is the eDNA distribution of nekton (fish, squid) species in the CAO?

As expected, the fish genome mapping and 12S amplicon methods identified the fish eDNA distributions in a comprehensive way. This eDNA study confirms that polar cod *Boreogadus saida* is the most frequently occurring fish species in the CAO, being represented in 50-60% of the samples. It was also found that the DNA of some commercial Atlantic predatory fish is more widespread in the CAO than expected. It is possible that these species find refuges at the Lomonosov Ridge and the Morys Jesup Rise, even if they in are deep (>1000 m). This concerns mainly Atlantic cod *Gadus morhua*, haddock *Melanogrammus aeglefinus* and beaked redfish *Sebastes mentella*, but also Greenland halibut *Reinhardtius hippoglossoides,* snailfish (Liparidae) and eelpouts (Zoarcidae). eDNA of Atlantic cod also had a broad distribution over the deep basins and was even detected in the Canada Basin. Since the work was performed with a highly sensitive method, the records could of course be DNA traces imported from non-CAO regions by ocean currents, while being stabilised by the prevailing low temperature (below 0 °C. However, the eDNA distribution of the fish preferentially at ridge areas are not random distributions, as well as the records for squid exactly in the area where the squid was recorded with the FishCam (see WP3). These results should be confirmed by confirmatory fish and squid catches in the future, and suggest that this benchmarking study can be used to identify areas where fish catches are more likely to be successful.

(2) What is the eDNA distribution of the major fish prey species in the CAO?

The zooplankton community composition reflected and confirmed what is already known about the zooplankton communities of the CAO (see WP4), but the eDNA methods used (COI marker mapping, COI amplicons) seemed to overlook important prey species for fish, especially amphipods (see WP5). The underlying reasons for this are probably that amphipods are believed to shed only low amounts of DNA to the environment. The use of eDNA to detect the organisms that shed low amount of eDNA in water needs calibration improvement. In general, the zooplankton eDNA distributions detected were not as explicit as those of the fish, i.e., there were no specific areas in the CAO where the eDNA of certain zooplankton (groups) occurred more than in other areas. This is probably because the pelagic zone of the CAO is a rather uniform environment with high connectivity, and with deep water even in ridge areas such as the Lomonosov and Gakkel Ridges. However, for biodiversity studies eDNA can be a crucial method for detecting unknown occurrences or high abundances of zooplankton species, especially concerning the microzooplankton.

(3) What is the eDNA distribution of the major fish predator species in the CAO?

Identifications of fish-eating marine mammals were extremely rare in the data. Only a few COI amplicons were detected in a few samples from narwhal *Monodon monoceros*, walrus *Odobenus rosmarus*, harp seal *Pagophilus groenlandicus*, and ringed seal *Pusa hispida*. These low occurrences were far too low for making eDNA distribution maps. The low occurrences of sequences in the samples is most probably related to the paucity of eDNA traces of these animals in the environment because they are so rare in the CAO.

(4) *How did the new bioinformatics pipelines work for the CAO?*

The emphasis of WP7 was on exploring the feasibility of using metagenomic and eDNA data for answering questions 1, 2, and 3. This required development of new methodology in bioinformatics. In this project, a methodology for quantitatively extracting the eDNA

distribution of marine invertebrates (zooplankton, squid) and fish from eDNA in metagenomic data sets was successfully developed, refined and benchmarked. For marine mammals the method would most probably work as well if these animals would have been more abundant in the CAO.

It was observed that that using fish genome assemblies was a useful approach for identifying DNA from fish found in the CAO to the species level. One limitation here is the availability of genome assemblies for species found in the Arctic Ocean. It does not appear that a chromosome-level assembly is needed for this type of analysis. For example, the contig/scaffold level assemblies obtained from short read data for both *Boreogadus saida* and *Arctogadus glacialis* were sufficiently contiguous for this analysis. Though it is important that the genome assembly process is handled carefully, especially with respect to contamination, it appears that additional draft genome assemblies would be quite useful here.

In general, the method seems to work well. Sufficient compute and storage resources are required in order to store the raw data and run the analyses. In addition, the databases built and used for these analyses need to be dealt with carefully in order to prevent misidentification and/or false positives. For example, in the initial genomes database built, one genome that was heavily contaminated with non-fish sequences (e.g., human, bacteria, viruses) was found. The sequencing strategy (e.g., read type depth etc) appeared to be sufficient.

7.7. Relevance of the WP7 data for fish stock modelling

The eDNA results obtained by WP7 could identify some areas where fish catches are more likely to be successful to gain more knowledge on fish distributions in the CAO. These areas are in the first place the Greenland side of the Lomonosov Ridge, the Morys Jesup Rise, and the Siberian side of the Amundsen Basin (including the adjacent Lomonosov Ridge). In the future, eDNA inventories will doubtlessly be an important tool that can augment and refine inferences and stock estimates based on acoustics and fish sampling. This does not least pertain to monitoring invasive or protruding species entering oceanic regimes undergoing change because of climate effects. From the WP7 study it can be concluded that fish sampling is crucial for calibrating indirect methods (acoustics, eDNA). There is a need for expanding and curating genomic databases for application in fish stock inventories with Arctic species.

7.8. Recommendations from WP7 for the JPSRM of the CAOFA

eDNA samples should be taken routinely. Standard methods for eDNA sampling, filter types, extractions, sequencing and bioinformatics should be developed for inter-compatibility of the results. Contamination from humans, and marine fish, squid and shellfish as human food on board or fish bait should be avoided. For bioinformatics analyses the open-source pipelines, including reference databases, designed at SLU can be used. Metagenomic sequencing is preferred since it gives quantitative results.

Appendix: MOSAiC Data Policy (13 pp.)

2019-09-19

MOSAiC Data Policy

The Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) is a collaborative, international project to address pressing scientific questions in the central Arctic. The project's success, and its ultimate impact on science and society, relies upon professional coordination and data sharing across the participants. A transparent Data Policy is essential to achieve MOSAiC science objectives, to facilitate collaboration, and to enable broad use and impact of the MOSAiC data legacy.

Executive Summary

This Data Policy regulates data management, access and release as well as authorship and acknowledgment. Signing this Data Policy is a pre-requisite for participation in MOSAiC field operations and being a member of the MOSAiC consortium.

Metadata Standards (for details see section 3)

Metadata shall make data findable and provide additional contextual information about measurement details, methods, relevance, lineage, quality, usage and access restrictions of the data. It shall allow coupling users, software and computing resources to the data. Hence, metadata must be machine-readable and interpretable as well as human understandable. Furthermore, metadata for each data set should follow the FAIR data principles in terms of fitness for purpose and fitness for re-use.

Data Ingest, Transfer, Storage and Archiving (for details see sections 5 and 6)

The MOSAiC Central Storage (MCS) aboard Polarstern is the basis for gathering data during the year of operation, offering near-real-time access and early processing of the data to the users underway. The land MCS provided by AWI is the central and reliable storage and working database of MOSAiC data within the AWI storage platforms.

Only MOSAiC consortium members with authentication/authorization will have access to the data prior to public release.

PANGAEA is the primary long-term archive for the MOSAiC data set and all primary data, with the exception of the subsequently mentioned cases, must be submitted to the PANGAEA data base for long-term archival. If this is not feasible due to the size of the data set or is not possible due to institutional data policies or commitments to other stakeholders, exceptions can be made if the data are stored in another long-term archive that provides unique and stable identifiers for the datasets and allows open online access to the data. These exceptions need to be documented in written agreements between the data provider and the MOSAiC Project Board and data manager.

Data Provision, Access and Sharing (for details see section 7)

Early access by the members of the MOSAiC consortium to the data is crucial for the successful collaboration within the consortium. Hence, all data must be made available to the consortium by the MCS as fast as possible. The following deadlines mark the latest points in time for transferring data to the MCS:

- All sensor data: Must be stored in the onboard MCS as fast as technically possible. Data that cannot be stored immediately in the on-board MCS have to be added as soon as possible or stored in the land MCS no later than 31 Jan 2021. Buoy data can be updated within one month

after the lifetime of the buoy if data are being collected beyond the end of the MOSAiC expedition.

- All fast analysis sample data: Must be stored on the land MCS no later than 31 Jan 2021.
- A primary subset of laboratory sample analysis data: Must be stored on the land MCS no later than 31 Jul 2021.
- Full collection of laboratory sample analysis data: Must be stored on the land MCS latest no later than 31 Jan 2022.

All MOSAiC raw and primary data are freely available to all MOSAiC consortium members as soon as they are stored in the on-board MCS or the land MCS.

For using data from the MCS for publications, the *data provider* or *data PI* **must be informed and offered collaboration** on the scientific analysis and must be offered co-authorship based on the principles described in section "Authorship and Acknowledgment" below. The *data provider* and/or *data PI* may object to the usage of data in a publication if that publication conflicts with his or her own publication strategy. Any such objection must be discussed and agreed upon in writing with the MOSAiC coordinator and data manager. The *data provider* and/or *data PI* may not object to the usage of data beyond the public release date.

Public Release of Data (for details see section 8).

MOSAiC data will be freely and publicly available on the open MCS or PANGAEA and/or alternate public archives on **1 Jan 2023**. From this date on there are no restrictions on data usage, but data users are encouraged to communicate with *data providers* or *data PIs* during early stages of all scientific analyses to ensure accurate usage and interpretation of data. The best practices on co-authorships described in the section "Authorship and Acknowledgment" below continue to apply.

Authorship and Acknowledgment (for details see section 9)

Generally, **co-authorship** on publications and other public documentation must be offered to those that have **made a substantial contribution** following the principles of good scientific practice. An inclusive co-authorship approach is encouraged.

Accordingly, co-authorship on publications and other public documentation must generally be offered to those that have made a substantial contribution to a) the intellectual conception or design of research; b) the acquisition, analysis, or interpretation of the data (i.e., including the *data provider* or *data PI*), or c) the drafting or significant revision of the work.

Lead authors have the ultimate decision authority and responsibility to identify and appropriately engage co-authors.

Contributors to the work that do not warrant co-authorship should be identified by name in the acknowledgments.

MOSAiC data must be acknowledged or referenced in publications and other public documentation, specifically including relevant digital object identifiers, data providers (if not co-authors), and funding agencies.

All publications and other public documentation using MOSAiC data must include a funding acknowledgment of MOSAiC in general in the following form:

"Data used in this manuscript was produced as part of the international Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC) with the tag MOSAiC20192020". Additionally, the Project ID given for specific expedition must be mentioned. For the Polarstern expedition this is AWI_PS122_00. Additional attributions like specific award/grant numbers might be added.

Data Publication (for details see section 10)

The publication of MOSAiC data via data journals and data archives is strongly encouraged and will be facilitated by the MOSAiC Project Board and Data Group. The MOSAiC Project Board will centrally organize one or more special issues in a data journal, with an appropriate period for submission. These special issues will allow for linking all MOSAiC data sets and help to make data standards and procedures easily citable.

Responsibilities

Data Group Speaker

Stephan Frickenhaus

Data Manager (primary contact)

Antonia Immerz (Antonia.Immerz@awi.de)

Data Group

Atmosphere: Peter von der Gathen, Matthew Shupe (CU/NOAA), Sara Morris (CU/NOAA)

Ice/Snow: Marcel Nicolaus, Martin Schneebeli (WSL-SLF), Julia Regnery

Eco, Bio-Sampling: Allison Fong, Pauline Snoeijs-Leijonmalm (Se)

BGC: Walter Geibert

Ocean: Ben Rabe, Julia Regnery

Airborne: Andreas Herber

Remote sensing: Thomas Krumpen, Suman Singha (DLR)

Modeling: Ralf Jaiser

PANGAEA & data publishing: Daniela Ransby, Stefanie Schumacher, Amelie Driemel ([info@pangaea.de)](mailto:info@pangaea.de)

Infrastructure Experts: Peter Gerchow, Angela Schäfer, Ingo Schewe, Mohammad Ajjan

Head of Data at AWI: Frank Oliver Glöckner

Head of Systems at AWI: Christian Schäfer-Neth

NSF Arctic Data Centre: Christopher Jones, Jesse Goldstein, Matt Jones

ARM: Giri Prakash

1. Objective

The purpose of this Data Policy is to codify the goals and principles of MOSAiC's research data life-cycle from production, documentation, sharing, usage and re-usage. This ensures that common procedures for data gathering, archiving and publication, as well as metadata and quality management are commonly implemented. By participating in the MOSAiC project, all members of the MOSAiC consortium agree to and comply with this Data Policy. By doing so, participants ensure that MOSAiC is a successful and resource-effective research project that also supports data accessibility, interoperability and re-usage following the FAIR data principles.

This policy aims to:
1. Ensure proper storage, backup and archiving of MOSAiC data in a central system.
2. Promote the visibility and accessibility of MOSAiC data for scientific and other applications.
3. Ensure the fair and equitable use of MOSAiC data and uphold the rights of individual scientists and institutions.

4. Enable the organized and timely analysis of the data.
5. Encourage the rapid publication and dissemination of scientific data, results and knowledge, to support the involvement of a broad user community.

2. Definitions

- **MOSAiC data**: Data collected aboard Polarstern, within the Central Floe Observatory, within the distributed network, and aboard Polar 5/6. This includes data from analyzed sample material and sample metadata and satellite data products.
- **Collaborating data**: Relevant data outside of MOSAiC data, brought to the MOSAiC consortium via the endorsement process (external aircraft data, re-supply vessel data, other coordinated activities). As defined by the endorsement, these data from collaborating partners are subject to the MOSAiC Data Policy.
- **External data**: Relevant data outside of the MOSAiC data and Collaborating data, but still of interest to the MOSAiC consortium and other users of MOSAiC data, including but not limited to operational model output, operational observations at other locations, etc. These data may be archived or cross-linked along with MOSAiC data at the discretion of the data provider but are not subject to the Data Policy and the provider is not entitled to the benefits of endorsement.
- **Data provider/PI**: All data streams must have a responsible party. The data provider is defined as the PI or institution that owns and/or operates an instrument, creates and analyzes samples, produces a model output, or otherwise produces a data set.
- **Consortium members**: Participants whose scientific activities are officially endorsed by the MOSAiC Science Board. Such participants are bound to the MOSAiC Data Policy and will have access to MOSAiC data as soon as they arrive at the MOSAiC Central Storage (MCS).
- **Public users**: Public users are those that use MOSAiC data or Collaborating data but are not part of the MOSAiC consortium.
- **Raw data**: Data directly produced by sensors, devices, or manual observation, prior to additional processing, calibration and quality assessment/control (never modified).
- **Primary data**: Processed data that modify a copy of the raw data, e.g., outliers removed, calibrated, quality controlled.
- **Value-added data/derived data product**: Products based on raw or primary data that may involve derivation of additional parameters or delayed-mode quality control using external data or post-use sensor calibration; model data or a combination with any external data, e.g., by data assimilation, visualization, classification, or clustering.
- **MOSAiC Central Storage (MCS)**: Connected central storage infrastructure that allows for the redistribution of data to consortium data users with authentication and authorization. Part of the MCS is aboard Polarstern for gathering and securing raw and/or primary data.
- **MOSAiC Standard operating procedures (MSOPs)**: MOSAiC teams specify procedures on how to handle devices, how to store samples, and how to process data. MSOPs are temporarily stored in the MCS. MSOPs document how data are processed from raw to primary and/or value-added data. They need to be published at the time the data are published in an open access format. When revised, MSOPs are subject to version control. MSOPs become, like data, open access and citable.
- **MOSAiC sensor and device registration**: Sensors and sampling devices are registered and managed centrally using the SensorWeb interface provided by AWI. The sensor registration is mandatory for controlling data streams through MCS and serve to augment data with metadata automatically. The combination of sensor registration and MSOPs will facilitate a high standard of quality management and documentation for referencing in publications.
- **MOSAiC Device ID (MDID)**: All sensors/instruments in MOSAiC have a unique ID and Uniform Resource Name (DeviceURN) in SensorWeb.

- **MOSAiC Sample ID (MSID)**: Physical samples or materials carrying physical or biological matter (e.g., filters) must have a unique ID.
- **MOSAiC Device Operation ID (MDOID)**: IDs registered in the Ship data system DShip, referring to coordinates and time. They can be recorded automatically, semi-automatically, or manually.

3. Metadata Standards

Metadata shall make data findable and provide additional contextual information about measurement details, methods, relevance, lineage, quality, usage and access restrictions of the data. It shall allow coupling users, software, and computing resources to the data. Hence, metadata must be machine-readable and interpretable as well as human-understandable. Furthermore, metadata for each data set should follow the FAIR data principles in terms of fitness for purpose and fitness for re-use. The metadata should be agreed on, listed, and explained within the MSOPs.

Specifically, within MOSAiC the following two general principles for providing metadata to MOSAiC datasets shall be endorsed:

- Metadata for sensors/devices must be registered in the SensorWeb. The derived DeviceURN from SensorWeb for each device should always be linked within the metadata for each data set ingested into the MCS as well as any derivate data to keep track of the available standardized meta data in SensorWeb.

- Specifically, all metadata necessary for archiving must be provided within the MCS at the moment data sets are ingested on board to ensure proper data sharing, findability, and re-usability during the expedition and later on. If this is not possible, e.g., due to technical limitations, all relevant data must be added latest until the public release date.

Recommendations for metadata and vocabularies

If further metadata are needed within the MSOPs we recommend using this collection of widely accepted metadata standards categorized by disciplines and communities to be adopted by MOSAiC sub teams.

Examples of standards are:

- **Oceanography, climatology, and modelling**
  - [CF (Climate and Forecast) Metadata Conventions:](#) The CF standard was framed as a standard for data written in netCDF format, with model-generated climate forecast data particularly in mind. However, it is equally applicable to observational datasets, and can be used to describe other formats. It is a standard for "use metadata" that aims both to distinguish quantities (such as physical description, units, and prior processing) and to locate the data in space and time.
  - [ISO 19115:](#) An internationally adopted schema for describing geographic information and services. It provides information about the identification, the extent, the quality, the spatial and temporal schema, spatial reference, and distribution of digital geographic data.
  - [ISO 19115-2:](#) Imagery and gridded data as an extension of [ISO 19115](#) defining the schema required for describing imagery and gridded data.
- **Biology**
  - [Ecological Metadata Language](#) (EML): A metadata specification that is used to document environmental data from almost any scientific domain, and includes sections for describing spatial, temporal, thematic, and taxonomic coverage of datasets. Current release: EML 2.1.1.

- o [Darwin Core:](#) A body of standards, including a glossary of terms (in other contexts these might be called properties, elements, fields, columns, attributes, or concepts) intended to facilitate the sharing of information about biological diversity by providing reference definitions, examples, and commentaries. Current Biodiversity Information Standards (TWDG) from October 2009.

  - o [MIxS: Minimum Information about any (x) Sequence:](#) The MIxS is a unified standard developed by the Genomic Standards Consortium (GSC) for reporting of minimum information about any (x) nucleotide sequence. It consists of MIGS, MIMS and MIMARKS standards and describes fourteen environments. MIGS, MIMS and MIMARKS share common mandatory core descriptors, differ in standard-specific elements and can be tailored to a particular environment by a subset of relevant environment-specific information components.

- **Provenance**
  - o [W3C Provenance Ontology](#) (PROV-O): The PROV-O ontology provides terms that support the documentation of the lineage of activities (like data processing), used and produced resources (like data), and the agents (like scientists) associated with the activity. The [DataONE ProvONE ontology](#) extends the PROV-O ontology to explicitly capture lineage information for scientific workflows, and statements about data inputs, processing scripts, and data outputs can be expressed inside of DataONE packaging documents (OAI-ORE resource maps).

All variables and parameters (measurement attributes) must be documented with an attribute name and attribute definition that provides a human-readable context for the measurement. For numeric data, attributes must include the units of measurement using SI unit definitions. Where non-SI units are used, a mapping to SI units must be provided that includes a) a unit name, b) a unit definition, c) a unit notation abbreviation, d) the unit's parent SI unit name, e) a multiplier to the parent SI unit. For numerical data without a unit (e.g., percent, count x per count y, etc.), the unit should be noted as "dimensionless". For non-numeric, categorical data, coded values must be defined in a code/definition list, or be defined by an external, controlled vocabulary term. We recommend the **NERC Vocabulary Standard**, since registry of MOSAIC Sensors and devices via SensorWeb follows this vocabulary. The NERC Vocabulary Server (NVS) web service provides access to controlled vocabularies via an international, actively-contributing research community https://www.bodc.ac.uk/resources/vocabularies/. Any deviations from this recommendation must be individually discussed with the MOSAiC data manager. In case a specific vocabulary is agreed on, a mapping between the NERC vocabulary term and the term used in the metadata must be provided by the requesting party.

Recommendation for Processing Levels
Processing levels of all data stored in the MCS or published in PANGAEA or other certified repositories should be stated in the metadata. In general, the levels raw, primary and value-added/derived should be used (see definition above). If other conventions or standards for data levels exist these should be referenced in the metadata. Processed data in PANGAEA and other certified repositories should include the information how they have been derived from raw data (provenance). Additionally, the information how to gain access the raw data should be provided.

4. Metadata Registries

The purpose of metadata registries is to assemble provenance meta information for the discovery, quality assessment, interlinking, and assembly of otherwise disconnected data.

ActionLog – Actions are registered in the DShip system on board. Sampling, regular station visits, etc. can be recorded with an App on a specific MOSAiC tablet. The recorded logs are uploaded to DShip by the data support team aboard.

Devices registry – Sensors and sampling devices are registered in SensorWeb by the PIs with support from the data support team (on board, but mainly before expedition start). Configuration changes are registered in the same system.

5. Data Ingest and Transfer

The MCS aboard Polarstern is the basis for gathering data along the year of operation, offering near-real-time access and early processing of the data to the users underway.

The land MCS provided by AWI is the central, reliable storage and working database of MOSAiC data within the AWI storage platforms. It will furthermore serve to distribute data after the expedition, also for data publication in other repositories.

Raw data obtained during the MOSAiC expedition shall be stored in the MCS on Polarstern. Any deviations from this rule must be individually agreed upon with the data manager. The raw data are transferred to the on-board MCS semi-automatically. Additional data can be submitted manually to MCS via mobile external hard drives in 'delayed mode' by scientific cruise participants.

For the data ingest into MCS, the Raw Data Ingest Framework provided by AWI (RDIF/AWI) will be used. For this, sensor registration in SensorWeb is mandatory, as is naming a responsible person for data transfer to the MCS. A data set template is to be described for RDIF, implying a DeviceURN from SensorWeb, a filename filter as regular expression (RegEx), file format descriptions and additional metadata for PANGAEA (see annex).

The transfer of the raw data after each leg to the land MCS at AWI is organized centrally by the AWI data support team. Data transfer to the land MCS will be performed by means of mobile data storage mediums (hard disks) hereby also maintaining user rights. Data is then made accessible adhering to the specified user rights of all MOSAiC members. Furthermore, raw data transferred to the land MCS will be automatically archived in a WORM (write once, read multiple) system at AWI.

Primary data produced aboard Polarstern during the expedition can also be transferred to the land MCS at AWI via the centralized data transfer. User rights defined on the data will be maintained accordingly. Publication of primary data sets in PANGAEA or other recommended repositories is the responsibility of each scientist. Data copies will be made accessible to the participating institutes via the land MCS at AWI.

6. Data Storage and Archiving

The land MCS will store the data and metadata records during and beyond the duration of the MOSAiC project. It will serve as a working database for the early handling and exchange of data within the MOSAiC consortium. As stated in section 2, only consortium members with authentication/authorization will have access to the data until public release (see section 7 and 8).

The land MCS will be in operation and accessible until all pre-registered data from the expedition, and the associated derived and analyzed data and metadata are permanently archived and published.

PANGAEA is the primary long-term archive for the MOSAiC data set and all primary data, with the exception of the subsequently mentioned cases, must be submitted to the PANGAEA data base for long-term archival. If this is not feasible due to the size of the data set or not possible due to institutional data policies or commitments to other stakeholders, exceptions can be made if the data are stored in another long-term archive that provides unique and stable identifiers for the datasets

and allows open online access to the data. These exceptions need to be documented in written agreements between the data provider and the MOSAiC Project Board and data manager.

Metadata of primary data sets published in PANGAEA are provided in a machine-readable format via the website of PANGAEA and are harvestable. The completeness of the metadata is the responsibility of the data PI. This option to harvest the meta data enhances the global visibility of MOSAiC data.

In PANGAEA, data files are archived together with metadata. Its content is distributed via web services to portals, search engines, and catalogs of libraries and publishers. Each data set includes a bibliographic citation and it is persistently identified using a Digital Object Identifier (DOI). Interlinkage of MOSAiC IDs (links to, e.g., SensorWeb, sample IDs, Device IDs, Grant IDs) is possible and allows the clear identification of data, samples, methods and associated data flows. For a more detailed sketch of PANGAEA workflows and options see the annex.

Datasets stored in other well-established, long-term archives, e.g., due to requirements by national funding bodies, should nevertheless be reported to the data manager and PANGAEA to ensure long-term, robust linkage with and documentation of all data that are stored externally to PANGAEA.

Molecular data (DNA and RNA data) must be archived within one of the repositories of the International Nucleotide Sequence Data Collaboration (INSDC, www.insdc.org) comprising of EMBL-EBI/ENA, GenBank and DDBJ).

In any case, each data set must have a clearly identified primary archive. Any exceptions from the rules stated here need to be agreed on between the data provider and the MOSAiC Project Board and data manager.

7. Data Provision and Sharing among the MOSAiC Consortium Members

Early access by the members of the MOSAiC consortium to the data is crucial for the successful collaboration within the consortium. Hence, all data must be made available to the consortium by the MCS as fast as possible. The following deadlines mark the latest points in time for transferring data to the MCS:

- All sensor data: Must be stored in the onboard MCS as fast as technically possible. Data that cannot be stored immediately in the on-board MCS have to be added as soon as possible or stored in the land MCS no later than 31 Jan 2021. Buoy data can be updated within one month after the lifetime of the buoy if data are being collected beyond the end of the MOSAiC expedition.
- All fast analysis sample data: Must be stored on the land MCS no later than 31 Jan 2021.
- A primary subset of laboratory sample analysis data: Must be stored on the land MCS no later than 31 Jul 2021.
- Full collection of laboratory sample analysis data: Must be stored on the land MCS latest no later than 31 Jan 2022.

All MOSAiC raw and primary data are freely available to all MOSAiC consortium members as soon as they are stored in the on-board MCS or the land MCS.

For using data from the MCS for publications, the *data provider* or *data PI* **must be informed and offered collaboration** on the scientific analysis and must be offered co-authorship based on the principles described in section "Authorship and Acknowledgment" below. The *data provider* and/or *data PI* may object to the usage of data in a publication if that publication conflicts with his or her own publication strategy. Any such objection must be discussed and agreed upon in writing with the MOSAiC coordinator and data manager. The *data provider* and/or *data PI* may not object to the usage of data beyond the public release date.

8. Public Release of MOSAiC Data

Good progress of a highly collaborative and interdisciplinary project like MOSAiC requires open availability of data to a wide user audience as early as possible. At the same time, it is important to acknowledge the substantial work that goes into collecting, quality controlling, formatting, documenting, and releasing scientific data. MOSAiC policies pertaining to data use and acknowledgment aim to balance these two principles.

Data access and usage policies evolve in time according to a staged process outlined here, and in all cases the most data-restrictive approach is described while an accelerated publication of data is acceptable.

MOSAiC data will be **freely and publicly available** on the open MCS or PANAGEA and/or alternate public archives on **1 Jan 2023**. From this date on there are no restrictions on data usage, but data users are encouraged to communicate with *data providers* or *data PIs* during early stages of all scientific analyses to ensure accurate usage and interpretation of data. The best practices on co-authorships described in section 9 "Authorship and Acknowledgment" continue to apply.

9. Authorship and Acknowledgment

Authorship. Generally, **co-authorship** on publications and other public documentation must be offered to those that have **made a substantial contribution** following the principles of good scientific practice. An inclusive co-authorship approach is encouraged.

Accordingly, co-authorship on publications and other public documentation must generally be offered to those that have made a substantial contribution to: a) the intellectual conception or design of research, b) the acquisition, analysis, or interpretation of the data (i.e., including the *data provider* or *data PI*), or c) the drafting or significant revision of the work. Co-authors should understand the content of the work, be accountable for at least a section of the work and approve of the final draft. Additional standard guidelines for deciding on co-authorship on publications can be found via numerous on-line resources, such as http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html or https://www.dfg.de/sites/flipbook/gwp/files/assets/basic-html/page85.html.

Lead authors have the ultimate decision authority and responsibility to identify and appropriately engage co-authors.

Contributors to the work that do not warrant co-authorship should be identified by name in the acknowledgments.

Authorship conflicts may be resolved by the MOSAiC Project Board, possibly taking into consideration advice from further experts in the research field.

Acknowledging data usage. MOSAiC data **must be acknowledged or referenced in publications and other public documentation**, specifically including relevant digital object identifiers (DOI, see Section 7), data providers (if not co-authors), and funding agencies. A data acknowledgment or reference should also specify where the data was obtained, according to individual journal policies. A suggested format for acknowledging each data stream includes:

"*[Data descriptor]* data (*[Author name et al. (PubYear)]*) was provided by *[data provider, PI, and or Institution]* with support from *[Funding agency or institution]*.

The data has then to be cited in the References, e.g., as follows:

"Nicolaus, Marcel (2018): Shipborne visual observations of Arctic sea ice during POLARSTERN cruise PS106. PANGAEA, doi:10.1594/PANGAEA.889264, In: Hutchings, Jennifer K (2018): Shipborne visual observations of Arctic sea ice. PANGAEA, doi:10.1594/PANGAEA.889209."

Acknowledging MOSAiC in general. All publications and other public documentation using MOSAiC data must include a funding acknowledgment of MOSAiC in general in the following form:

"Data used in this manuscript was produced as part of the international Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC) with the tag MOSAiC20192020". Additionally, the Project ID given for specific expedition must be mentioned. For the Polarstern expedition this is AWI_PS122_00. Additional attributions like specific award/grant numbers might be added.

Citing Research Platforms. All scientific and data publications must cite the article concerning the respective research platform:

"Polarstern: Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung. (2017). Polar Research and Supply Vessel POLARSTERN Operated by the Alfred-Wegener-Institute. Journal of large-scale research facilities, 3, A119. http://dx.doi.org/10.17815/jlsrf-3-163"

"Polar5 and Polar6: Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung. (2016). Polar aircraft Polar5 and Polar6 operated by the Alfred Wegener Institute. Journal of large-scale research facilities, 2, A87. http://dx.doi.org/10.17815/jlsrf-2-153"

10. Data Publication

Clear, consistent documentation of MOSAiC data will help to support a strong and lasting MOSAiC data legacy, promote the broad and appropriate use of MOSAiC data including the citation of data, and ensure proper acknowledgment of data creators. This documentation is particularly important for a large, inter-disciplinary, and international project like MOSAiC, which involves many disparate sources and providers of data. To this end, the publication of MOSAiC data via data journals and data archives is strongly encouraged and will be facilitated by the MOSAiC Project Board and Data Group.

- Data publication can take multiple forms such as data journals or data/metadata archives (potentially certified by WDS/CoreTrust). Data publications follow the FAIR data principles. The ultimate goals for data publication are to provide a clear description of the metadata and data, the specific instruments and measurements that created the data, the quality control procedures, the manner in which the data were processed, any embedded data dependencies (on other data sets), and any other special conditions or considerations for the data. To assist in data tracking and awarding of credit, it is important that data sets are given a digital object identifier (DOI). Additionally, associated data files, metadata description documents, and processing scripts and instruments should receive a persistent identifier (PID), which links to the datasets.
- Authorship on data publications should follow similar policies to authorship on scientific publications and must include those participants that have made substantial contributions to collecting the data, processing the data, and documenting the data (see Section 9). Each data publication needs a contact person and principle investigator (PI) who is familiar with and responsible for the scientific evaluation. This is especially relevant for "automated" measurements, where often the cruise scientist is chosen as PI, but was not involved in the data evaluation.
- The MOSAiC Project Board will centrally organize one or more special issues in a data journal, with an appropriate period for submission. These special issues will allow for linking MOSAiC data sets and help to make data standards and procedures easily citable. Each special issue will likely have an introductory manuscript that provides the context for the rest of the special issue. When organizing the special issues, the coordinator will specify a short list of recommendations for the information that should be specifically included in data publications. This process might involve specific MOSAiC formatting that will support consistency across the different publications.

- External Data: When used in a publication in the MOSAiC context, i.e., in combination with MOSAiC data, external data should be published in an appropriate open access data repository that also provides DOIs or at least persistently resolvable IDs.
- Synthesis Data: MOSAiC data may serve as a basis for synthesis data products, i.e., data from MOSAiC in combination with already published data or model data. Synthesis data should be published in the same manner as MOSAiC data. PIs working on synthesis data and related publications are encouraged to ensure that data from other sources becoming part of synthesis data are published.

11. Amendments

Variances
Any modifications to this policy that are needed on a case-by-case basis, i.e., conflicting requirements from a funding agency, must be endorsed by the MOSAiC Project Board.

Dispute resolution
Disputes on the Data Policy should be solved primarily by the involved individuals or MOSAiC team leaders. If resolution at this level is not possible the MOSAiC Project Coordination will act as a mediator in the conflict. If resolution cannot be achieved with the mediation of the Project Coordination, the MOSAiC Project Board will be engaged to resolve the dispute.
In case, the MOSAiC Project Board is not able to resolve the dispute amicably it will be referred to the competent German state court. German law under exclusion of its conflict of law regulation and under exclusion of the Convention on the International Sale of Goods (CISG) will be applicable.

MOSAiC Consortium
The term "MOSAiC Consortium" does not refer to a legal entity or institution. MOSAiC Consortium defines a scientific collaboration of many persons contributing scientific work to the project. Consequently, the term "Official Member" refers to the fact that the person signing the data policy will respect the Consortiums Data Policy and that he/she is registered for book keeping on a formal basis, and for realizing the technical basis of data sharing.

Signature

Name	
Institute	
e-Mail	

Hereby I declare that I fully consent to the MOSAiC Data Policy and become a registered MOSAiC Consortium Member.

Date, Signature

12. Annex

Requirements for MOSAiC Sample IDs (MSID)

Physical samples or materials carrying physical or biological matter (e.g., filters) must have a unique ID. Also, certain measurements and data products, such as photographs for instance must obtain a unique ID.

Creation of unique sample IDs is to be managed within the scientific teams.

The association with the device and its operation in which the sample was obtained must be documented. Therefore, the respective DeviceURN and DeviceOperation ID must always be related to a sample ID. This is achieved by annotating sampling log sheets enlisting sample-IDs with the DeviceURNs from SensorWeb of the involved devices and the DSHIP-DeviceOperation IDs in which the device was deployed. Storing the sampling log sheets in the respective directory of the MCS which reflects this structure exactly makes the metadata clear to the data user.

PANGAEA - sketch of workflows/options and metadata

Datasets in PANGAEA may be archived as stand-alone publications of data (e.g., https://doi.org/10.1594/PANGAEA.753658) or as supplements to an article (e.g., https://doi.org/10.1594/PANGAEA.846130). Data can be submitted to and published in PANGAEA with access restrictions in place for a predefined period (until article publication, or during an embargo period). Metadata must be submitted together with the data (minimal requirements are dataset Author(s), PI, dataset title, MOSAiC ID(s), related institute(s) or publication(s)). Any documentation (e.g., MOSAiC Standard operating procedures, MSOPs) helping to understand the data can and should be linked to the dataset(s). If no persistent link to the documents can be provided, PANGAEA can archive the documents permanently alongside the data.

The granularity of the data is up to the author(s) of the dataset. Lower-granularity datasets can be combined in a time-series collection dataset as in https://doi.org/10.1594/PANGAEA.873032. During submission (https://www.pangaea.de/submit/), the connection with MOSAiC has to be clearly stated in the Label Field of the Data Submission. The MOSAiC Project ID (see Acknowledging MOSAiC in general, section 9) must be given in the Data Submission description. The MOSAiC Device ID(s) should also be provided. Within the data table, parameters (table header) should be submitted with full names and units. Data submitted in the form of videos, photos, geoTIFF, shape files, netCDF, sgy, etc. will be archived as is (e.g., https://doi.org/10.1594/PANGAEA.865445). More information on data submission can be found in https://wiki.pangaea.de/wiki/Data_submission.

If a published dataset needs to be updated, PANGAEA will upload a new version of this dataset, with new documentation and complete metadata (clearly providing information on the changes between the versions). Both versions can be linked but will have their own permanent DOI.

MOSAiC Grant IDs

MOSAiC grant-IDs are provided centrally by the MOSAiC science board via the MOSAiC Project Board. Grant-IDs are parse-able for analyzing citations within the Acknowledgments in papers referring to MOSAiC, see Acknowledging MOSAiC in general, section 9. Additional grant IDs from funding agencies might exist.

Appendix: SAS 2021 Research Data Management Policy (4 pp.)

Background

The Swedish Polar Research Secretariat (SPRS) recognises that high quality research data are valuable products of field work. Data sharing, with fair attribution, is a corner stone for scientific collaboration, not least for the evolvement of interdisciplinary science. Data preservation with open access is fundamental for the legacy of any research activity. Facilitating reuse and re-purposing of research data adds long-term value of the data to scientific research, industry, and society at large.

Building on the data statements of the International Council for Science (ICSU)[1] and the International Arctic Scientific Committee (IASC)[2], SPRS works towards ethical open data publication pertaining supported projects, programmes and research infrastructure.

SPRS seeks to promote high standards and best practice for management of research data. Data originating from research projects on-board Icebreaker Oden during the SAS 2021 expedition are to be published openly, with proper attribution to the data creator, and with minimum delay, unless otherwise agreed upon.

Research data and metadata

Research data refers to any information necessary to validate and reproduce the results of research. This could be e.g. field notes, primary data files, images, or audio-visual materials.

Metadata refers to structured information about the data. This information includes technical information (such as file formats) as well as descriptions of provenance and

---

[1] http://www.icsu-wds.org/services/data-sharing-principles

[2] https://iasc.info/images/data/IASC_data_statement.pdf

context (purpose of study, timeframe, sampling locations, equipment used etc.). The metadata should be descriptive and detailed enough to enable independent interpretation and re-use of the data.

Published research data and metadata should be accurate, complete, identifiable, and openly accessible, and strive to meet the FAIR principles[3]. The dataset together with accompanying metadata should be archived securely and safely in appropriate formats to ensure long-term usability. The data publication should also be given a unique persistent identifier enabling citation, versioning, and proper attribution.

Open access to data and metadata

Data collected during the SAS 2021 expedition can be divided into three categories with corresponding open access policies.

1) *Data originating from installations on Oden funded and/or operated by SPRS:*

   SPRS is owner of data. Metadata and data are to be made available with free, unrestricted, complete access, without charge and with minimum delay in the Swedish National Data Service (SND) data repository[4] after completion of the cruise.

2) *Data collected during the expedition within the scope of a separately funded research project that requires specific installations or activities outside the premises of paragraph 1:*

   Either the organisation of the principal investigator (PI) or the PI is owner of data, depending on applicable regulations or agreements. Metadata and data are to be made available with free, unrestricted, complete access, and without charge within the framework of current legislation. Metadata are to be submitted to the SNDs data repository or other suitable data repository with minimum delay after completion of the cruise. Access to data can be restricted by the PI for a maximum of two years (period of moratorium). If the funder of

---

[3] https://www.force11.org/fairprinciples

[4] http://snd.gu.se

the research project has stipulated a shorter period of moratorium this has precedence.

3) *Data collected by SPRS in agreement with third party:*

   The agreement should stipulate open access policy and data ownership. SPRS should always strive towards full open access with minimum delay.

The principal investigator is responsible for:

- informing themselves on SPRS research data policy, including which open access policy applies to their data.

- developing appropriate and well documented procedures for data collection, processing and use.

- prior to the expedition, submitting requested information regarding how data will be collected, managed, shared, and openly published.

- ensuring safe and secure storage of data and metadata during and after the cruise. The PI determines which information needs to be retained to support the authenticity of any research results.

- openly publishing data and metadata according to the principles outlined in this document in a format facilitating re-use and re-purposing.

- submitting complete metadata for publication in the SNDs data repository in association with other metadata and data from the SAS 2021 expedition.

- if data is not submitted for publication in SNDs data repository, ensuring that the repository selected for data publication meets requirements of long-term preservation. Links to the published dataset should be submitted to SPRS for publication in SNDs data repository.

- following all agreements with regard to deadlines, publication, data-sharing, moratorium and/or confidentiality.

The Swedish Polar Research Secretariat is responsible for:

- providing PI with advice regarding data management and publication.

- If requested supporting PI in primary publication of metadata and data in the SNDs data repository.

- watching over the long-term preservation of data and metadata published in the SNDs data repository and if needed taking relevant actions regarding data custodianship.

- publishing links to data on polar.se and/or other relevant communication channels controlled by SPRS.