Framework Service Contract EEA/IDM/15/026/LOT 1 for Services supporting the European Environment Agency's (EEA) implementation of cross-cutting activities for coordination of the in situ component of the Copernicus Programme Services

ARCTIC IN SITU DATA AVAILABILITY

Erik Buch, Marianne Sloth Madsen, Jun She, Martin Stendel, Ole Krarup Leth, Ann Mari Fjæraa, Mikael Rattenborg

Issue: 2.1
Date: 09/12/2019

[Document Release]

	Name(s)	Affiliation
Coordinated by:	Erik Buch	EuroGOOS
Contributions:	Marianne Sloth Madsen Jun She Martin Stendel Ole Krarup Leth Ann Mari Fjæraa Mikael Rattenborg Erik Buch	DMI DMI DMI DMI NILU Consultant EuroGOOS
Approval:	Henrik Steen Andersen	European Environment Agency

Prepared for:	European Environment Agency (EEA)
Represented by: (Project Manager)	Henrik Steen Andersen
Contract No.	EEA/IDM/15/026/LOT1

[Change Record]

Version	Date	Changes
1.0	31/05/2019	First draft release
1.1	26/06/2019	Updated first draft release
1.2	31/08/2029	Updated first draft release
2.0	08/11/2019	Updated version after evaluation meeting with representatives from Copernicus Services and Satellite component
2.1	9/12/2019	Updated version after review by Matthew Fry

List of content

Executive Summary ***5***

1. Introduction ***7***

2. Requirements for Arctic In Situ Data ***9***

2.1. Meteorology **9**
2.1.1 Introduction to C3S 9
2.1.2 C3S Requirements for Arctic in situ data 9

2.2. Atmospheric composition **11**
2.2.1 Introduction to CAMS 11
2.2.2 Use of Arctic in situ data in CAMS 12
2.2.2.1. The overall CAMS requirements 12
2.2.2.2. The CAMS validation activity (CAMS-84) requirements 14

2.3 Ocean **14**
2.3.1 Introduction to CMEMS 14
2.3.2 Requirements of in situ observations in the Arctic 15
2.3.2.1 Arctic MFC (ARC MFC) 15
2.3.2.2 TACs 17
2.3.2.3 Oceanographic data requirements for C3S 20

2.4. Cryosphere **22**

2.5 Copernicus Space Component **22**

3. Overview of existing Arctic in situ data ***25***

3.1. Meteorology **25**
3.1.1. Freely available data 25
3.1.2. Data with restricted availability 32

3.2. Atmospheric composition **33**
3.2.1. Freely available data 33
3.2.2. Data with restricted availability 42

3.3. Ocean **43**
3.3.1 Data with open access 46
3.3.1.1 Water temperature and salinity 46
3.3.1.2 Currents 51
3.3.1.3 Sea level and waves 53
3.3.1.4 Sea ice 55
3.3.1.5 Biogeochemistry 56
3.3.1.6 Data identified by INTAROS 60
3.3.1.7 Data usage in ARC MFC and satellite TACs 63
3.3.15 Data with restricted availability 71

3.4. Cryosphere **72**
3.4.1. Freely available data 72
3.4.1.1 Ice sheets (Greenland) 73
3.4.1.2 Glaciers 73

3.4.1.3 Permafrost 74
3.4.1.4 Snow 74
3.4.2. Data with restricted availability 76

4. Gap analysis 78

4.1. Meteorology 78
4.1.1. Observations do not exist 78
4.1.2. Observations exist but data not freely available 78
4.1.3. Technology gaps 79
4.1.4. Sustainability gaps 79

4.2. Atmospheric composition 79
4.2.1. Observations do not exist 79
4.2.2. Observations exists but data not freely available 80
4.2.3. Technology gaps 80
4.2.4. Sustainability gaps 80

4.3. Ocean 80
4.3.1 Data gaps and adequacy analysis 81
4.3.3 Technology gaps 84
4.3.4 Sustainability gaps 85

4.4. Cryosphere 86
4.4.1. Observations do not exist 86
4.4.2. Observations exists but data not freely available 86
4.4.3. Technology gaps 86
4.4.4. Sustainability gaps 86

5. Discussion and conclusions 87

General conclusions 88

Specific thematic domain related conclusions 89

6. Next step and recommendations 92

Immediate actions 93

List of Acronyms 95

References 100

Executive Summary

The Copernicus community's requirements for environmental in situ data from Arctic region has been collected together with information on the existence of such data – freely available or restricted. Comparing the two sets of information reveals severe gaps in:

- The present Arctic Observing System – especially the central Arctic is undersampled
- Timely availability and quality of existing observations
- Availability of data from non-European countries
- Fit-for-purpose of observation technology
- Sufficient data management structures at data producer level
- Sustainability of existing observing system – many rely on time limited research funds

As a practical way forward, Copernicus is recommended to consider the feasibility of the following short-term actions that need to be implemented in a coordinated and collaborative manner:

- Continue to define and document Copernicus specific requirements to an Arctic in situ observing system – attention should especially be on:
    - Resolution in space and time
    - Data quality improvement
    - Meta data improvements
- Establish formal links to intergovernmental bodies such as SAON, WMO, IOC and GEO to secure that Copernicus' requirements for a sustained and integrated observing system are articulated and taken into account
- Continue to setup and leverage international cooperation arrangements between the EU and non-EU countries with an Arctic interest, e.g. Canada, South Korea, Japan, and the USA;
- Liaise with Horizon Europe to promote that:
    - Arctic relevant observing technology and data communication development is included in future research calls – focus could be on multipurpose and autonomous observing platforms
    - Research projects are requested to secure a free exchange of data along the FAIR principle using existing European data management infrastructures
- Pursue innovative cost-effective technological solutions for Arctic observations securing continuous NRT data flow from this harsh environment also during wintertime
- Initiate data rescue activities composed of but not limited to the following components:
    - Continuous support of projects like the C3S inventory effort, enhanced data collection, homogenization and mining;
    - Maintain and further develop centralised data portals for the individual thematic domains;
    - Start a task force focussing on unlocking existing data not currently available to Copernicus. The effort could include support to organisations without a proper data management structure, support to implementation of proper data quality control procedures;
    - Improve access to Russian data sources for all thematic themes in cooperation with relevant Russian authorities.
- Work with national authorities to
    - Secure sustainable funding of a fit-for-purpose Integrated Arctic Observing System

- Support initiatives toward real-time delivery of data
- Increase involvement of indigenous people in data collection

- Initiate the development of a European airborne survey programme as a contribution to fill the gap from Operation IceBridge

(Detailed recommendations are given in Chapter 6)

1. Introduction

The Arctic is a region very sensitive to environmental changes. There is a very close interrelation and delicate balance between atmosphere, terrestrial, cryosphere, sea ice and ocean, especially in relation to solar energy retainment, the radiation budget and the hydrological cycle. This has a great impact on physical, chemical and biological processes in the area.

Due to the hostile environment, there is a great lack of basic observations in the Arctic that can support scientific understanding of key processes. Most of the existing data are collected via time limited research projects. This lack of process knowledge is reflected for example in substantial errors in forecasting models – operational as well as climate.

It is foreseen that monitoring of the Arctic region will rely heavily on satellite observations supplemented by more traditional in situ platforms. The ocean community in particular will continue to use several other platforms such as ships, profiling floaters, gliders, moorings, AUV's etc. to monitor the interior of the Arctic Ocean. In addition, earth observation satellites rely heavily on precise in situ observations for calibration of satellite sensors and validation of satellite measurements.

The Copernicus Services and Space Component have at different occasions raised strong concerns about the timely availability of sufficient relevant in situ data from the Arctic region.

Consequently, the Copernicus In Situ Coordination Activity led by the EEA has initiated a project focusing on clarifying to what extent the necessary in situ data (near real-time as well as delayed quality controlled) are available to:

- Maximize the exploitation of present and future Copernicus Sentinels
- Produce and validate products from the Copernicus Services, in particular CMEMS, C3S, and CAMS

The objective of the project has been to provide an analysis of:

- The requirements for meteorological and ocean (incl. sea ice and cryosphere) in situ data in the Arctic region by Copernicus Services and Space Component. Land observations were not included in the project mandate.
- The existence and availability of the required data incl. identification of conditions for access to restricted data (payment, limitation in use, etc.)
- Any gaps identified in the observation system.

The work has been structured to address these three objectives:

Task 1. Overview of requirements for Arctic in situ data

- Dialog with CAMS, CMEMS, C3S, ESA and EUMETSAT representatives including a questionnaire survey
- Extraction of available information from CIS^2 and OSCAR
- Extraction of available information from H2020 projects, primarily the Integrated Arctic observation system (INTAROS), EU-Polarnet

Task 2. Overview of existing in situ data used by or relevant for Copernicus

- Open and freely available data

- Restricted data
- Sources of information
    - WMO – Global Atmosphere Watch (GAW) and Global Cryosphere Watch (GCW)
    - INSTAC
    - EMODnet incl. its Arctic checkpoint
      SAON and AMAP
    - INTAROS and iCUPE
    - JCOMMOPS
    - C3S 311a Inventory Activity
    - European Monitoring and Evaluation Programme (EMEP)
    - EU-Polarnet
    - IASOA
    - AERONET, NDACC, WOUDC, NILU, IAGOS, EEA

3. <u>**Gap analysis**</u>

- Based on the overview generated in task 1 and 2 a gap analysis has been performed

Geographical focus

In this study the boundaries of the Arctic are defined by the Arctic Monitoring and Assessment Programme (AMAP) and shown in Figure 1.1

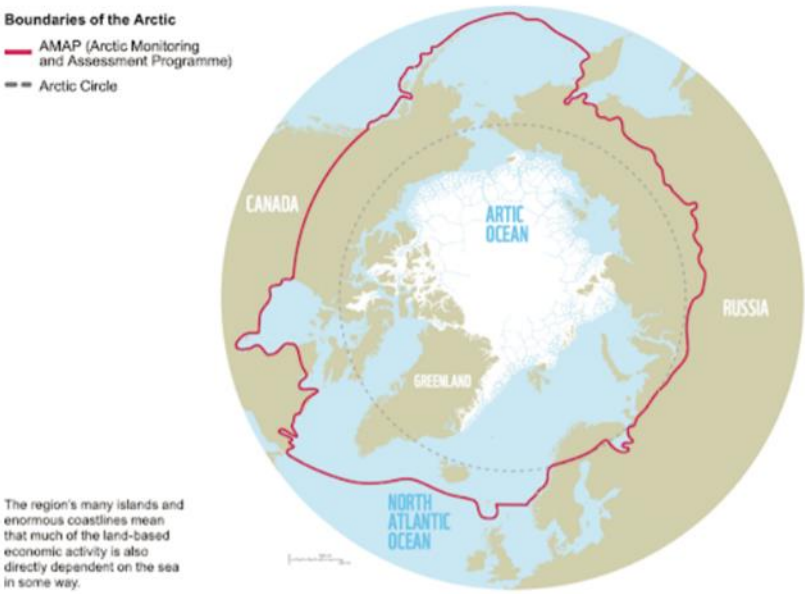

Figure 1.1 Arctic Boundaries as defined by AMAP

2. Requirements for Arctic In Situ Data

The requirements for Arctic in situ data within C3S, CAMS, CMEMS, ESA and EUMETSAT have been collected by distributing a questionnaire to the relevant services. The questionnaire asks for specific requirements and gaps related to the availability of Arctic in situ data needed by the services. Details of the replies will be presented and analysed in this chapter. Additional relevant sources of information are:

- WMO requirements are defined in the OSCAR database (https://www.wmo-sat.info/oscar/requirements)
- The Copernicus in situ coordination information system maintains a portal with data products and requirements (https://cis2.eea.europa.eu). A large number of essential variables are listed with information about data policy, specific requirements and compliance level.
- The Global Climate Observing System (GCOS) regularly assesses the status of global climate observations of the atmosphere, land and ocean and produces guidance for its improvement. The ECV factsheets provide overviews of the individual ECVs including products, requirements, and data sources https://gcos.wmo.int/en/essential-climate-variables/ecv-factsheets

2.1. Meteorology

2.1.1 Introduction to C3S

The Copernicus Climate Change Service (C3S) focuses on monitoring, analysing and predicting climate change with the aim of providing authoritative information about past, present and future climate in Europe as well as globally. C3S is implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF).

In C3S, in situ data are used to produce climate reanalyses and seasonal forecasts and for independent validation of the products. Furthermore, a number of C3S projects provide Climate Data Records for Essential Climate Variables derived from satellite observations, and these require in situ data for validation. Another main objective of C3S is to provide reliable access to the above-mentioned products through the Climate Data Store (CDS). The CDS will also provide tools and guidance to enable end-users to analyse and visualize their data as needed for their specific purpose. The establishment of the CDS is still ongoing and data sets are continuously being added.

2.1.2 C3S Requirements for Arctic in situ data

The project has collected input on in situ data requirements from the C3S global reanalysis (ERA5), the C3S Arctic Regional Reanalysis and the C3S seasonal forecasts. Information from the C3S European Regional Reanalysis project has not yet been received, but as only a small part of their domain is located in the Arctic this is not considered a major issue.

We also collected requirements from C3S 312b: Essential Climate Variable products derived from observations. We received responses from Lot1 (atmospheric physics), Lot2 (Atmospheric composition), Lot3 (Ocean) and Lot4 (Land hydrology and cryosphere). Lot5 (land and biosphere) replied that they did not rely on in situ data.

An overview of main requirements for meteorological information is given in Table 2.1.1 C3S requirements for ocean and cryosphere data are included in the respective sections.

Table 2.1.1 C3S requirements for meteorological data from the Arctic

Product/purpose	Requirements	Current data	Comments
ERA5 global reanalysis (data assimilation and independent validation)	Surface and mean sea level pressure 2m temperature and humidity 10 m marine wind Upper air temperature wind and humidity Accuracy according to state-of-the-art instrumentation.	Mainly from GTS https://software.ecmwf.int/wiki/display/CKB/ERA5+data+documentation#ERA5datadocumentation-Observations	Timeliness: within 6 hours Sub-daily time resolution Data are assimilated into the reanalysis products.
Arctic Regional Analysis/ Data assimilation and validation	Temperature, wind and humidity profiles For validation, other parameters such as: precipitation, cloud cover, visibility and radiation parameters	Mainly GTS Some data by request from partners	As accurate as possible. 3-hourly data Timeliness: 1.5 h for data assimilation, 'within reason' for validation.
Seasonal forecasts Data assimilation for initialization; model development and validation	Meteorological and land surface observations including upper air data (for NWP initialization) Surface radiation (representative locations including Central Arctic)	Real-time data from GTS Data availability over Arctic Ocean is far from adequate	Access to original (non-gridded) data is preferred QC is part of data assimilation
Global satellite products for Precipitation, water vapour, cloud properties, Earth Radiation budget, surface radiation budget. In situ data used for validation/verification as well as correction/ calibration.	Daily mean values of precipitation and radiation in the Arctic	Data are freely available and downloaded regularly Providers: NOAA (daily precipitation + radiation) BSRN: World Radiation Monitoring Center - Baseline Surface Radiation Network (BSRN)	Best possible accuracy and spatial coverage. QC: quarterly checks for consistency Timeliness: 1-2 months delay

In general, the C3S services request a good spatial coverage - as good as possible - and would benefit a lot from having more data from data sparse regions. For C3S, it is important to have long homogeneous time series, and it is therefore very important to have good documentation of instrumentation and methodologies. Comprehensive metadata are also requested in order to assess the quality of the data.

Project-specific requirements for meteorological variables are:

C3S ERA5 Reanalysis
In situ data are especially important for the Arctic, as not all satellite data can be used in the polar region. Surface and upper-air meteorological variables are required. Access to data not freely available would be highly beneficial. (Long) records for sea ice cover observations are highly beneficial for verification of ingested sea-ice-cover products near complicated coastlines.

C3S Arctic Regional Reanalysis
The regional reanalysis has a very high resolution of 2.5 km and therefore needs in situ observations with a very good spatial coverage. Therefore, a large effort has been made to get access to existing data which are not freely available. Any additional data in data sparse regions would be beneficial, and especially more SYNOP stations over Greenland are requested

C3S Seasonal Forecasts
For seasonal forecasts, spatial coverage is an issue. Data availability over the Arctic Ocean is far from adequate; most of what is available is not operational and may not be sustained without further support. A wide range of data (ocean, ice, land, atmosphere, stratosphere, including ozone) would be useful for validation.

C3S Essential Climate Variable products derived from Observations (Lot 1: Atmosphere)
The project is providing global satellite products for five ECVs: Precipitation, water vapour, cloud properties and radiation budget (Earth and surface). In situ reference data are needed for validation and verification, in particular for the Arctic region where (re-analysis) models have difficulties. The radiation products need to be calibrated with Baseline Reference Surface Radiation (BSRN) observations as they have difficulties in snow-covered regions (with very high and varying albedo). Daily mean values from in situ observations of precipitation and radiation is needed. Better spatial coverage would be beneficial, and long-term availability is a key issue.

2.2. Atmospheric composition

2.2.1 Introduction to CAMS

The Copernicus Atmosphere Monitoring Service (CAMS) provides consistent and quality-controlled information related to air pollution and health, solar energy, greenhouse gases and climate forcing, everywhere in the world. CAMS is implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF).

To provide and further develop the CAMS portfolio, ECMWF works with many service providers around Europe. By doing so, CAMS combine the expertise and infrastructure that exist in Europe to provide a range of services that are unequalled by any other organisation in the world.

To acquire all the observations that are needed to produce the CAMS services, ECMWF collaborates with the European Space Agency (ESA) and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) as well as many other organisations providing satellite and in situ observations.

The core of the CAMS service is providing access to data and tools related to atmospheric monitoring. CAMS also support users by providing quality assurance information and advice on how to use and interpret data.

An extensive catalogue of data visualisations in the form of maps and charts with data from a variety of sources is provided.

The CAMS data catalogue itself includes daily analysis and forecasts of reactive gases, aerosols, European air quality, ozone layer, CO2 and fire monitoring, available as regional products, global products, supplementary products and emission products.

The present report focuses solely on the regional Arctic data and not the full spectra of CAMS products and services.

2.2.2 Use of Arctic in situ data in CAMS

A questionnaire on Arctic data requirements was distributed to the CAMS consortium. The aim of the questionnaire is to assess to what extent the necessary Arctic in situ data are available to initialize, assimilate and validate the CAMS products.

The project has collected input on in situ data requirements from the Evaluation and Quality Control element of CAMS (CAMS-84 contract), in addition to one response covering all CAMS services in general. The identified requirements from the responses in the questionnaires are listed below.

2.2.2.1. The overall CAMS requirements

Some of today's most important environmental concerns relate to the composition of the atmosphere. The increasing concentration of greenhouse gases and the cooling effect of aerosol are prominent drivers of a changing climate, but the extent of their impact is often still uncertain. At the Earth's surface, aerosols, ozone and other reactive gases such as nitrogen dioxide determine the quality of the air around us, affecting human health and life expectancy and the health of ecosystems. Ozone distribution in the stratosphere influences the amount of ultraviolet radiation reaching the Earth's surface. Dust, sand, smoke and volcanic aerosols affect the safe operation of transport systems and the availability of power from solar generation, the formation of clouds and rainfall, and the remote sensing by satellite of land, ocean and atmosphere. To address these environmental concerns there is a need for data and processed information.

CAMS has been developed to exactly meet these needs, aiming at supporting policymakers, business and citizens with enhanced atmospheric environmental information. The Service consolidates many years of preparatory research and development and delivers the following operational services:

- Daily production of real-time analyses and forecasts of global atmospheric composition
- Reanalyses providing consistent multi-annual global datasets of atmospheric composition with a stable model/assimilation system
- Daily production of real-time European air quality analyses and forecasts with a multi-model ensemble system

- Reanalyses providing consistent annual datasets of European air quality with a frozen model/assimilation system, supporting in particular policy applications
- Products to support policy users, adding value to "raw" data products in order to deliver information products in a form adapted to policy applications and policy-relevant work
- Solar and UV radiation products supporting the planning, monitoring, and efficiency improvements of solar energy production and providing quantitative information on UV irradiance for downstream applications related to health and ecosystems
- Greenhouse gas surface flux inversions for CO_2, CH_4 and N_2O, allowing the monitoring of the evolution in time of these fluxes
- Climate forcing from aerosols and long-lived (CO_2, CH_4) and shorter-lived (stratospheric and tropospheric ozone) agents
- Anthropogenic and natural emissions for the global and European domains and global emissions from wildfires and biomass burning

There is no specific CAMS line for the Arctic region alone, but as the atmosphere does not contain any boundaries, all CAMS services can be applied to the Arctic region. A few examples that cover Arctic areas are the monitoring of the ozone layer, transport of smoke from extensive wild fires in North America and Siberia, and deposition of pollutants.

CAMS can benefit from high-accuracy in situ data in many forms, but most importantly for evaluation and quality control of the various CAMS service lines. For the accuracy requirement for in situ, it is particularly important to have an accurate estimate of the observation error as this is essential for good quality control of the CAMS products.

The timeliness requirement of the in-situ observations is not quite as critical as it is for satellite observations that are used in the CAMS forecasting system, but it still matters. If observations arrive within days they can be used to validate near-real-time events. If data arrives later, even years later, they are still of great value and may be used to produce reanalyses and to run test experiments for historic periods.

The time resolution requirements range from hourly to monthly. Data with high temporal frequency (e.g. hourly) is most useful, but daily and monthly means can still be of great value. The spatial resolution requirement is that for data to be most useful in verification, they should be from fixed sites with well-documented metadata; data from moving sites such as ships are more difficult to use. Unfortunately, not many Arctic stations fulfil this requirement.

Sometimes there is a need to reject significant volumes of data because the file is not properly formatted, or the wrong value is used as missing data indicator, etc. CAMS would probably prefer to receive problematic data earlier and undertake its own QC rather than wait for months. Ideally there would be multiple processing levels - a level-0 for immediate data with little/no QC, a level-1 that arrives slightly later with some basic QC and a level-2 that arrives significantly later with maximal QC.

The CAMS service is interested in all sort of atmospheric composition observations both from in situ observations and ground-based remote sensing. CAMS is also very interested in observations of wet and dry deposition fluxes for a wide range of atmospheric constituents. Observations of pollen are also of interest to CAMS.

2.2.2.2. The CAMS validation activity (CAMS-84) requirements

The CAMS validation activity (CAMS-84) is providing validation reports on a 3-monthly basis for the real time products of CAMS (forecasts, analysis) as well as real time verification results in the form of direct comparisons of CAMS concentrations and available real time independent observations.

In particular, CAMS-84 aims to test if pollution episodes (transport) and background pollution levels are described well by the CAMS global system. The variable requirements are ozone, CO, aerosol (AOD, PM). Other quantities of use are aerosol specification, hydrocarbons, and greenhouse gases. The combination of column and surface observation is of importance (e.g. ozone sonde, in situ ozone, AOD+PM, FTIR+ surface CO).

The accuracy requirement is not very critical, but accuracy should be sufficient to quantify the variability in time of the trace species.

The spatial resolution is currently limited due to the limited number of stations (IASOA) and the large area concerned. For an assessment of the individual episodes of transport of pollution to the Arctic a larger number of ground stations would be helpful. The time resolution requirement is 3-hourly or better. The timeliness requirement is within 1 month after sensing for the validation reports, and within 1 day after sensing for the real-time monitoring. The delivery time requirement is continuous or daily to be preferred.

2.3 Ocean

2.3.1 Introduction to CMEMS

The Copernicus Marine Environment Monitoring Service (CMEMS) provides regular and systematic information on the state of the physical ocean at global and regional level. There are four main areas of benefit covered by the service: maritime safety, coastal and marine environment, marine resources, and seasonal weather forecasting and climate activities.

CMEMS is based on a production structure covering two layers:

- processing of space and in situ observations and delivery of derived products, through Thematic Assembly Centres (TACs) where in situ data is managed by the In Situ TAC (INSTAC);
- processing of models, for forecasts, hindcast and reanalysis, fed by products derived from space and in situ observations (to be provided by the TACs), through Monitoring and Forecasting Centres (MFCs), structured according to regional domains (6 European regional seas) and global ocean.

In agreement with the Delegated Regulation on Copernicus data and information policy, the products delivered by CMEMS are open and free of charge, and compliant with European regulations, such as INSPIRE. The products are accessible through a European one-stop-shop, the CMEMS web-portal, which includes a structured information catalogue monitored to ensure that it complies with its operational obligations to users.

2.3.2 Requirements of in situ observations in the Arctic

Results in this section are a blending of the present survey on ocean in situ data and CMEMS in situ requirement report (Mercator Ocean et al. 2018). The survey was designed and taken by DMI, to identify current usage of in situ data in CMEMS TACs and ARC MFC, and also their requirements. Answers have been received from Pierre Prandi (SEA LEVEL TAC), Jacob L. Høyer (SST TAC), Rasmus T. Tonboe and Gorm Dybkjær (Ocean Sea Ice TAC, through an interview), Arne Melsom and Laurent Bertino (ARC MFC).

2.3.2.1 Arctic MFC (ARC MFC)

In situ observations are needed in ARC MFC for calibrating the models, validating model products and reducing errors in the model fields via data assimilation. A description of general requirements on in situ observations for Arctic forecast and reanalysis is given in Table 2.3.1. There are in total 10 different applications, ranging from offline Calibration/Validation (Cal/Val) of physical ocean products to offline data assimilation in biogeochemical reanalysis. All activities need in situ data with Arctic coverage.

For Cal/Val purpose, a limited number of stations with high frequency sampling is an ideal solution. For generating Quality Identification Document (QUID) report and historic reanalysis, only offline observations are needed. For Near Real Time (NRT) analysis and forecast product validation, NRT data are needed. Applications in interim scales, i.e., 1-12 months before present, observations have to be available in corresponding time. Considering scientific and service purposes, priority areas of observations are given to Nordic Seas, Straits and important shipping areas e.g. Cape Farewell, Kara Sea and Northern Sea routes. For biogeochemical product validation, coastal shelf seas are important monitoring areas. Freshwater inputs are also important for models.

Table 2.3.1 General requirements on in situ data for Arctic forecast and reanalysis: PHY – Physics, BGC – Biogeochemical, DA – Data Assimilation, RAN - Reanalysis

Purposes/applications	Coverage	Resolution	Frequency	Timeliness	Priorities
PHY_Offline_CalVal	Arctic	Low	High	Offline	Nordic Seas, Straits, Northern sea route
PHY_NRT_CalVal	Arctic	Low	High	NRT+offline	
BGC_Offline_CalVal	Arctic	Low	High	Offline	Coastal-shelf seas
BGC_NRT_CalVal	Arctic	Low	High	NRT+offline	
PHY_Forecast_DA	Arctic	Min_res.	Any	NRT	Areas with high model uncertainties; coastal-shelf seas
BGC_Forecast_DA	Arctic	Min_res.	Any	NRT	
PHY_Interim_RAN_DA	Arctic	Min_res.	Any	1month	
PHY_Historic_RAN_DA	Arctic	Min_res.	Any	Offline	
BGC_Interim_RAN_DA	Arctic	Any	Any	1month	
BGC_Historic_RAN_DA	Arctic	Any	Any	Offline	

For forecast and reanalysis applications, requirements on sampling frequency are not critical. However, in order to obtain positive impacts, spatial resolution of in situ data needs to reach a certain level. It is still an open question on how to determine the minimum resolution needed. One way is to use impact studies such as Observing system experiment (OSE) or Observing System Simulation Experiment (OSSE). One example can be seen from the T/S profile observing system. Xie et al., 2017 have documented a reduction of bias and Root Mean Square (RMS) errors of up to 50% in both temperature and salinity in the whole Arctic by assimilating T/S in a reanalysis for a period of 1992-2016. In total 114952 water temperature profiles and 109575 salinity profiles were assimilated, corresponding to about 12 profiles per day or 4600 profiles per year. The spatial distribution of the stations is shown in Figure 2.3.1. The reduction of the temperature biases is gone 1 year after the assimilation stops. Due to such long temporal impacts of in situ data, even 12 profiles per day in the Arctic Ocean can generate significant positive impacts. Hence 12 T/S profiles per day in the Arctic can be regarded as the minimum spatial resolution required for data assimilation.

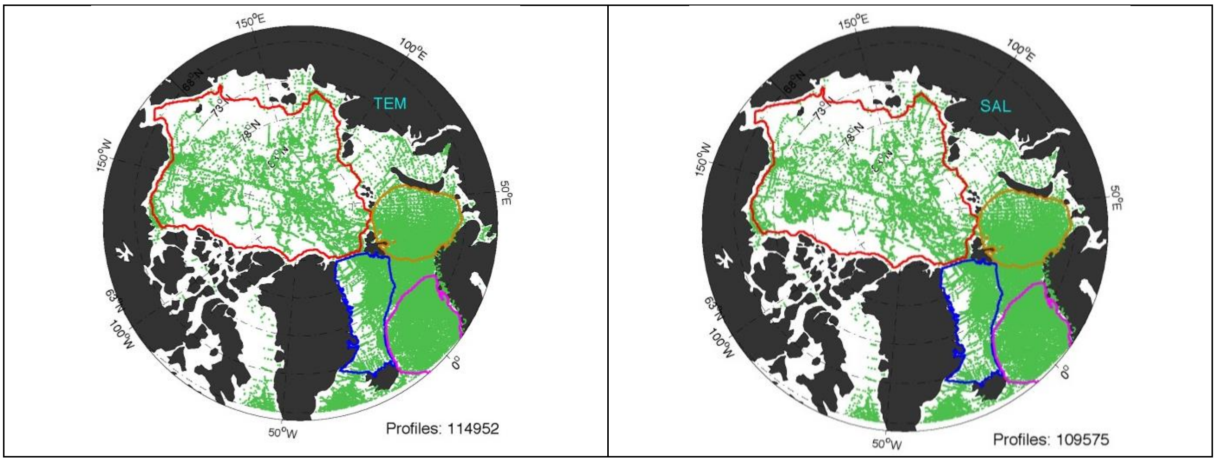

Figure 2.3.1 All assimilated Temperature (left) and Salinity (right) profiles in the ARC MFC Physical reanalysis in the whole period 1992-2016.

Table 2.3.2 Overview of the impact and priority assessment of in situ data in the ARC MFC (Bertino, 2018). Impact: "Real" = Has been documented. "Ongoing" = is being documented. "Ind.": Indirect through calibration of remote sensing products. "Science": Data is used for process studies. "None": Data not available or insufficient. Priority A = urgent, B= Second priority, C: Not a priority for the time being. (Source: Laurent Bertino, 2018)

OBS	T/S	SST	Surf. Currents	Current Prof.	Ice Drift	Nutrients	Waves	SL	Ice Thick	Snow
Impact	Real	Ind.	Science	Science	Ind.	Ongoing	None	Ind.	Ind.	None
Priority	A	B	B	B	C	A	A	B	A	A

For physical parameters, an assessment of in situ data impact and priority is given in Table 2.3.2. The impact lines show that some data sources may be more important to the Arctic MFC as an indirect resource for calibrating satellite data or because of their contributions to the theories that support models and remote sensing products. The Priority line however reflects the degree of urgency of intensifying the observation network. There is only one data under priority "C" (ice drift) because in the present situation the ice drift products from satellite data have good spatial coverage, but this does not mean that the in situ time series should be interrupted. The data needs of process studies for model developments and remote sensing products will be discussed later.

More detailed requirements on spatial resolution, temporal resolution, timeliness and data management were also provided by ARC MFC:

Spatial resolution.

- The spatiotemporal sampling distance for T/S profile should be 200 km in the ocean daily (equivalent to 50 profiles per day), higher resolution in key areas (the Fram Strait and the Barents Sea opening for example).
- The spatial resolution for biogeochemical variables should be even higher.
- The spacing of 200 km is necessary for drifting buoys as well.

Temporal resolution

- It would be good if all platforms (Argo, Ice Tethered Profilers - ITPs) could report data at least once a week in order to catch the shoaling of the mixed layer and the onset of the primary production bloom.

Timeliness

- For NRT assimilation, data can be assimilated with a latency up to 10 days old, but not older.
- Interim applications need data in the interim time scale.

Requirements related to data management

- The access to data should not require the installation of specific software (like OceanDataView or other) that are not practical for NRT runs. The users should be allowed to extract data automatically by date and allowed to download the whole time-series on the command line.
- Quality controlled observations are needed. CMEMS requires MFCs to update the reanalysis regularly until 1-month behind NRT, so the QC'ed data would be most useful if it comes within 1 month and 1 week from present time.

2.3.2.2 TACs

In TACs in situ data are needed mainly for calibrating and validating satellite algorithms and products. In situ observations normally need higher quality, more comprehensive metadata and sometimes higher spatiotemporal resolution than for model validation. However, requirements for spatial coverage are not as critical as in the model applications. All data are offline. For interim reprocessed data, in situ data should be delivered within one month.

Sea Level TAC (SL TAC):

Parameters: Tide gauge observations are needed to validate sea level gridded product (Level 4). Data from drifters and profiling floats are needed to derive currents for validating absolute geostrophic current derived from altimetry. More information about this dataset can be found in the PUM and QUID of the In situ TAC (http://cmems-resources.cls.fr/documents/QUID/CMEMS-INS-QUID-013-044.pdf and http://cmems-resources.cls.fr/documents/PUM/CMEMS-INS-PUM-013-044.pdf). The T/S profiles from Argo floats are needed to derive the steric sea level, for validating the altimeter total sea level.

Quality: in situ sea level data should be provided with standardized format, similar to (Permanent Service for Mean Sea Level) PSMSL data, and also with quality flags.

Ocean Colour TAC (OC TAC):

Parameters and quality: For the development and qualification of ocean colour primary variables of satellite sensors, high quality in situ measurements accompanied with their uncertainties are needed. These data are referenced as Fiducial Reference Measurements (FRM) by the ocean colour community.

The FRM observations include:

- Automated in-water radiometry from moored Fixed-Point Observatories
- Automated above-water radiometry from Tower Fixed-Point Observatories
- In-water radiometry from profiles acquired during dedicated research cruises
- Inherent Optical Properties (IOPs) from profiles acquired during dedicated research cruises

Biogeochemical measurements (i.e. Chlorophyll, Total Suspended Matter - TSM) and IOPs from samples acquired during dedicated research cruises. The in situ chlorophyll data are obtained through fluorometric or spectrophotometric techniques and from HPLC (high-performance liquid chromatography) measurements. In situ chlorophyll fluorescence methods are not deemed of adequate quality for product development or qualification, unless periodically calibrated with HPLC sample measurements. Primary production observations are needed for Cal/Val.

These observations are acquired and analysed following stringent measurements and data reduction protocols as defined by the space agencies and ocean colour community.

Sea Surface Temperature TAC (SST TAC):
***Parameters*:** within the SST TAC, surface drifting buoys, Argo floats and moored buoys are currently used to support the validation and monitoring of the operational and reprocessed L3 and L4 SST products. The main variable used is the water temperature. Sea ice surface temperature observations (Ice Surface Velocity Drifter - ISVP buoys) from drifting buoys for the validation of the Arctic SST/IST products are also required.

Metadata: in addition to measurement values and standard variables (coordinates, time, etc.), for all platforms and data streams additional information related to the in situ acquisition system is required to allow selection of data with confidence for validation. It is considered useful to include metadata detailing the network and/or project to which the platform belongs, the instrument/sensor model type, the deployment date, the sampling frequency, the depth of the sensor (mainly for surface platforms) and the drag loss. Moreover, the upcoming High Resolution (HR) SST drifters from Trusted projects (http://www.cls-telemetry.com/copernicus-trusted-project/) will report additional variables that must be included in the data files, as e.g. the statistical moments and/or percentiles of the subsample distribution, the complete set of high frequency data and the near surface water pressure which can be used to estimate the depth.

The satellite community therefore wishes metadata related to the in situ acquisition system to be propagated throughout the netCDF data files delivered by CMEMS. This includes:

- Network or project to which the platform belongs: for instance, the upcoming set of FRM buoys funded by Copernicus and EUMETSAT
- Instrument model types: for instance, to select buoys equipped with HR SST sensors
- Deployment date: to assess the age of the buoy and possible issues
- Averaging period: the period over which individual samples of a sensor are averaged to construct the reported value
- Sampling frequency: the frequency at which a sensor samples (for the subsamples used to construct the reported value)

- Depth of the sensor: the depth of the sensor for surface platforms (drifters)
- Drag loss: as estimated from data analysis or reported by the platform itself
- Calibration information: as provided by the buoy manufacturer
- Manufacturer

These platform/instrument metadata may be provided for instance in global attributes or in netCDF groups for files containing measurements from different platform or instruments.

In terms of parameters, the upcoming HR SST drifters from the Trusted projects will report additional variables that must be included in the data files, for instance:

- The statistical moments and/or percentiles of the subsample distribution
- The complete high frequency data (that will be broadcasted in some cases)
- The manufacturer
- The near surface water pressure sensor (can be used to estimate depth) In terms of temporal resolution, the highest temporal resolution available is required, in particular for moored buoys. HR SST drifters from Trusted project will also occasionally report the full samples (instead of a 5 minutes average every hour), these high-resolution data shall be reported in CMEMS product.

***Quality*:** generally, there is a need for a dedicated “surface” dataset with good quality control and selection of “good” surface data.

Timeliness: the requirement for satellite validation is to have the in situ measurements for a given day available the following day, . Moreover, the daily files shall be updated every day as long as new measurements are collected (meaning that if some buoy measurements are received 3 days after acquisition, these newly collected measurements shall be appended to the corresponding daily file). Users can then choose to wait (to have more data or better data) or not depending on their application.

Ocean Sea Ice TAC (OSI TAC):
***Parameters*:** Ship measurements from ice breakers are important. Ship-, land-, or aircraft-based in situ observations of icebergs are very useful. Intergovernmental Maritime Organization (IMO) Polar Code requirements have prompted ship operators, including a larger number of tour vessels, to start making routine sea ice observations to support the mandated risk assessment tools. Such systematic measurements could be used in e.g. validation of sea ice concentration, and lead to improvements in OSI TAC algorithms and products.

Specifically, the necessary in situ observation parameters are: temperature profiles in the ocean, ice, and snow. Such profiles are available from IMB (Ice Mass Balance) buoys and also ITPs. These two buoy types were used more extensively in the period around the International Polar Year (2007-2009), but since that period such buoys have become scarce. In addition, Sea Ice Drift (SID) data are also needed.

***Quality*:** OSI TAC needs improved access to ice drifting buoys. The measurement error for SID should be less than 100 m. Currently there are only a few, which are used to validate products as well as improving and developing detection algorithms. For the algorithms there is a special focus on ice thickness and snow on ice. When measuring ice drift, the most important in situ information comes from having precise Global Positioning System (GPS) positions. ARGO or Iridium positions are not precise enough. Gaining access to polar observations in NRT can be difficult. Arctic and Antarctic buoys are not coordinated through e.g. an overarching program, and this makes it challenging to access NRT data.

Wave TAC:
Parameters: The wave information used for validation depends on the dataset:

- For Synthetic Aperture Radar (SAR) data, which involves directional spectral wave observations, significant wave height, mean wave period, wave steepness, peak wave lengths and wave directions can be derived; only buoys measuring and providing the 5 mandatory parameters for the 2D spectrum reconstruction are considered.
- For the altimeter data, which involves the total significant wave height observations, all buoys measuring this parameter can be considered.

Data characteristics: hourly sampling frequency with a decent spatial coverage in ice-free waters (both in coastal and open ocean waters) are needed. Considering that the modelling and forecasting wave-ice interaction processes is still on-going, integrated in situ observations are needed (both wind and wave parameters) for R&D purpose.

2.3.2.3 Oceanographic data requirements for C3S

Requirements for oceanographic data were also identified from the C3S survey as described in Section 2.1.2. C3S requirements for ocean and sea-ice data are summarized in Table 2.3.3.

Table 2.3.3 Sea ice requirements from C3S. GLOSS - Global Sea Level Observing System; CLIVAR - Climate Variability and Predictability

Product/purpose	Requirements	Current data	Comments
ERA5 global reanalysis	Sea ice data for verification of the ERA5 sea-ice product.		
Seasonal forecasts	Sea ice thickness Snow on ice Melt ponds	GTS for real time data Multiple providers	Prefer access to original (non-gridded data)
Sea ice thickness/ validation	Sea ice free-board, sea-ice thickness/draft Snow depth and density Snow grain size and wetness (preferably)	Airborne surveys, ice tethered buoys and moorings.	Timeliness is not critical
SST based on satellite data (C3S)/ Validation	In situ SST data from CMEMS in situ TAC	Sensors can be dropped on or trapped in the ice. This can cause large biases but is not indicated in the metadata.	Timeliness not critical
Sea ice concentration products/ Validation	In situ data with a resolution of 1x1 km Data should cover large areas and range of ice types. Data are required throughout the year	Ice charts from the National Ice Center (NIC) are currently used for. In situ data would be valuable as ice charts have limitations.	Timeliness is not critical
Sea level/ Validation	Not specified	Tide gauge data GLOSS/CliVar & PSMSL	Timeliness is not critical.

For C3S seasonal forecasts, the required ocean variables include mixed layer temperature, salinity, sea-ice thickness, sub-surface ocean profiles (good coverage), melt ponds and snow on ice. Most of these have already been covered by CMEMS. Only melt ponds were not mentioned in CMEMS. These data are needed for data assimilation, model development and validation.

In the future, Arctic sea-ice cover is likely to become more variable as the average thickness in summer approaches zero. Good quality data to initialize seasonal forecast models is likely to become increasingly important.

For C3S in situ ocean variables are also needed for validation of Essential Climate Variables derived from observations. Timeliness is in general not critical as the primary purpose is quality control and validation. Quality control is usually done by the provider of the data. Long-term availability is crucial to assess the stability of the sea ice products.

Sea-ice thickness:

In order to be useful in a C3S context:

- Airborne surveys should have a regional coverage of significantly more than 100 km.
- Mooring data need the length of a full Arctic winter.

A significant spatial and temporal lack of sea ice thickness data for validation is communicated, especially for the Russian Arctic. Two reasons are mentioned:

- Drifting ice buoys (out of Russian Arctic)
- Russian Arctic underrepresented in terms of airborne observation programs.

For validation of sea ice thickness, snow depth and density are needed but it is beneficial to have snow grain size and wetness as well, as these properties affect the radar altimeter signal. It is also relevant to have buoy measurements of snow surface height changes to constrain snow depth on sea ice.

It is noted that the quality of sea ice thickness information will improve with the evolutions of satellite sensor technology. An example is the potential CRISTAL – Copernicus polaR Ice, Snow and Topography Altimeter, which dual-frequency altimeter concept may eliminate the need for auxiliary data by measuring freeboard and snow depth directly.

Sea ice concentration

For sea ice concentration (SIC) and edge products, ice charts are currently used for validation. The uncertainty could be reduced, if in situ data could be used. The validation data should cover large areas and multiple ice types. A resolution of 1x1 km would be needed for validation, maybe higher resolution in the future due to higher resolution instruments. In situ data are required throughout the year. Access to data within 1 month is required for routine validation reporting.

SST

For SST, consistency in quality between current and historical data and long–term stability of data sources is needed. For sea level, more stations with GPS measurements are required.

2.4. Cryosphere

According to the Essential Climate Variable (ECV) fact sheets, the ECVs for the cryosphere are:

- Ice sheets and ice shelves: Surface elevation change, ice velocity, ice mass change, grounding line location and thickness
- Glaciers: Glacier area, Glacier elevation change, glacier mass change
- Permafrost: Thermal state of permafrost, active layer thickness
- Snow: Area covered by snow, snow depth, snow water equivalent

The C3S requirements for cryosphere in situ data in the Arctic are summarized in Table 2.4.1.

Table 2.4.1 C3S requirements for cryosphere data

Product/purpose	Requirements	Current data	Comments
ERA5 global reanalysis/ Data assimilation and model validation	Snow depth	Data from GTS and national reports	
Arctic Reanalysis/ Data assimilation and validation	Snow depth at least once a day, every hour for validation		
Seasonal forecasts/ Data assimilation and validation	Snow on ice Melt ponds	GTS for real time data Multiple providers	Prefer access to original (non-gridded data)
Satellite observations of Greenland ice sheet surface elevation change (SEC) and gravimetric mass balance (GMB)	In situ measurements of surface properties in the snow (from firn/ice coring) Repeated air-borne laser scanning at annual intervals	Data from NASA Operation IceBridge Data from ESA campaigns	The cutting of Operation IceBridge increases the need for European collection of validation data
Greenland Ice Sheet Velocity/ Validation	GPS ice velocity; accuracy: <0.1 m/year; point data on Greenland (fast) outlet glaciers and in the (slow) interior.	Data obtained through request from field partners and from Cryosat	Irregular timeliness Time resolution 6d to 1 year contempo-raneous with product resolution; Delivery time: annual

The C3S projects generally require snow data for data assimilation and validation**.** In situ data with higher resolution (in time and space) are requested and foreseen to be even more important to validate future products which use higher resolution instruments. In situ data are required for validation of surface elevation changes of the Greenland Ice Sheet – these have been provided by Operation IceBridge and a gap is foreseen in the future

2.5 Copernicus Space Component

The Copernicus Space Component, operated by EUMETSAT and ESA require in situ data for the following Calibration/Validation activities:

- Validation against **precise in situ measurement – Fiducial reference Measurements (FRM) (few points but precise and well characterized).** FRM measurements are used for validation

(or correction) of the instrument calibration, as well as for validation of primary geophysical products.
- Validation against **in situ measurements** (more points and less precise) for validation of geophysical products.

For some instruments dedicated ground-based transponders are use, and these typically reside in easily accessible non-arctic areas, an exception being Svalbard where transponder infrastructure is hosted.

The requirements for FRM-level in situ data are primarily addressed by the agencies themselves, as these are critical for the mission return on investment. However, for the broader in situ measurements, the requirements are overlapping with the requirements of the Copernicus services.

Following the questionnaire, ESA has provided a preliminary list of requirements for Arctic in situ data, shown in Table 2.4.1, covering ocean and cryosphere. Included are requirements for the Sentinel-1, -2 and -3 missions. The table does not include requirements for transponder infrastructure to support instrument calibration as well as in situ data requirements for future Sentinel missions.

Note: Specific timeliness requirements for operational Sentinel Calibration/Validation activities have not yet been documented, but typically in-situ for operational validation of satellite products are need within 48 hours after measurement.

Table2.4.1. ESA requirements

Activity	Parameters	Sentinel Missions	Relevant In situ sensors
Ocean validation			
	sea level (sea surface height)	S-3	Tide gauge, Moving vessel profiler (MVP), Gliders, CTD, Argo buoys, Drones
	ocean wave	S-1, S-3	tide gauge, buoys
	sea surface gradients/current	S-1, S-3	Moving vessel profiler (MVP), Gliders, CTD, ADCP, drifting buoys, HF radar, drones
	sea surface temperature	S-3	SST radiometers, MVP, gliders, CTD
Sea-Ice validation			
	sea ice thickness	S-3	boat and airborne campaigns (i.e mosaic), submarine data, drifting buoys
	sea ice free board	S-3	boat and airborne campaigns (i.e mosaic), submarine data, drifting buoys
	snow depth (+temperature and salinity)	S-3	boat and airborne campaigns (i.e mosaic),
	sea ice drift	S-1, S-1, S-2	boat and airborne campaigns (i.e mosaic),
	Iceberg drift, size and thickness	S-1, S-2, S-3	

Land-Ice validation			
	ice sheet spectral albedo, broad band albedo	S-3	
	ice sheet snow grain size distribution	S-3, S-2	
	Ice sheet surface temperature	S-3	
	ice sheet surface elevation	S-3	Campaigns
	ice sheet surface velocity	S-1, S-2	Campaigns
	ice sheet/shelf calving front location	S-1	Campaigns
	Permafrost active layer seasonal subsidence	S-1	

3. Overview of existing Arctic in situ data

3.1. Meteorology

3.1.1. Freely available data

Real-time atmospheric observations from a multitude of sources are being shared globally via the Global Telecommunication System (GTS) of the WMO Information System (WIS). GTS is implemented and operated by National Meteorological Services and International Organizations, such as ECMWF and EUMETSAT, to ensure that all WMO Members have timely and reliable access to all meteorological and related data, forecasts and alerts.

GTS data are exchanged according to WMO Resolution 40 (Cg. XII) and Resolution 25 (Cg XIII) for meteorological, and hydrological data respectively. Annex 1 to Resolution 40 defines a set of essential data that each member shall exchange without charge and with no conditions on use. Each member should provide as many data as possible, but at least those that will assist in defining the state of the atmosphere at least on a scale of the order of 200 km horizontal resolution and six to 12 hours in time. The essential data thus include, 6-hourly SYNOP data, all marine in situ observations, all available aircraft reports as well as all data from upper air sounding networks.

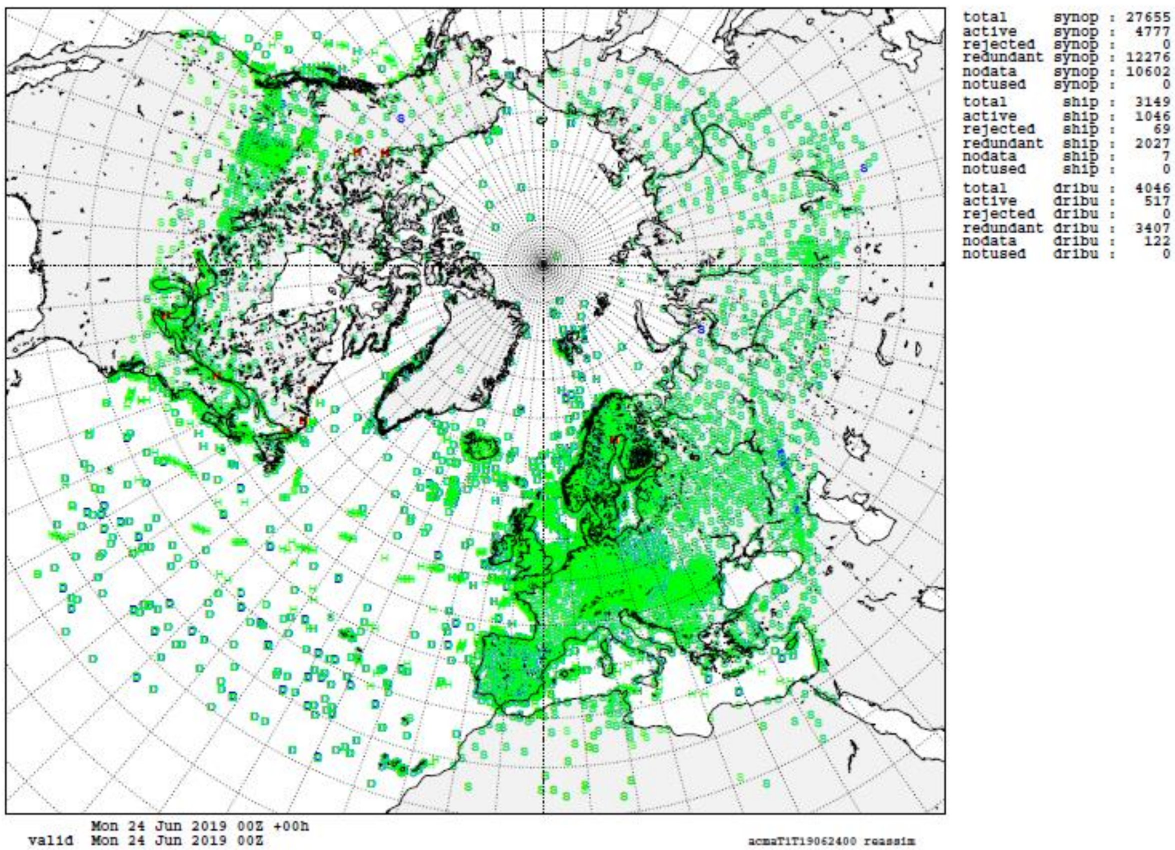

Figure 3.1.1 Arctic surface observations from SYNOP stations, ships and drifting buoys. The map shows data used in the Arctic Reanalysis project; these are mainly received from GTS.

Figure 3.1.1 shows a typical distribution of surface observations by SYNOP stations, ships and drifting buoys. While the density over land is reasonable, there are large gaps over the Arctic oceans and the inner Arctic. In particular, there is a lack of observations over northern and eastern Greenland, east of Svalbard, east of the Lena delta, around the Hudson Bay, over the Labrador Sea as well as over large parts of the coastal areas of Nunavut and Yukon.

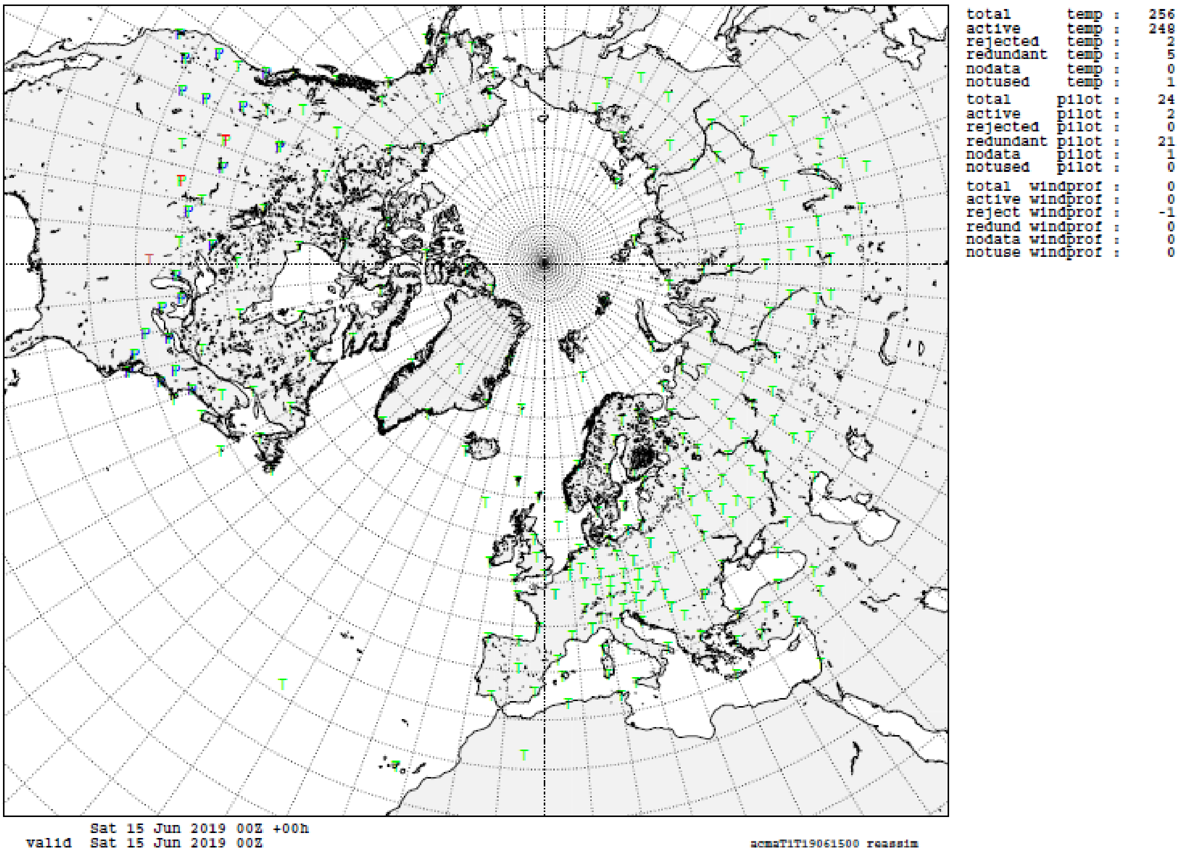

Figure 3.1.2 Upper air measurements as received from GTS

Figure 3.1.2 shows a typical distribution of upper air measurements. There are very few observations north of 70°N. Note that there are no radiosonde observations at all in the inner Arctic, only 5 on Greenland and virtually none over the whole Canadian Arctic. After the end of the cold war, several (expensive) weather ships have been closed down (in particular “C” southeast of Greenland and “M” east of Iceland). We note that there is no upper air information from radiosondes available as well. Note that radiosonde ascents are devised to be conducted twice daily (00 UTC and 12 UTC), however, numerous stations only send up one radiosonde per day, most often at 00 UTC.

Figures 3.1.3 and 3.1.4 show typical distributions of upper air observations obtained from aircrafts for two different dates. Flight observations from the WMO AMDAR Observing system apply to WMO standards and are made available from GTS. Information is also gained from aircraft reports (AIREPS) which include air temperature, wind speed and direction as well as aircraft position. AIREPS are used through data exchange arrangements with the International Civil Aviation Organization (ICAO).

Aircraft data are quite variable in time and also in space, depending on the weather conditions. This is most clearly visible over Greenland and to a lesser extent also over Siberia. Again, there is almost no data in the inner Arctic, although there occasionally are observations from flights across the Arctic from Europe to Alaska.

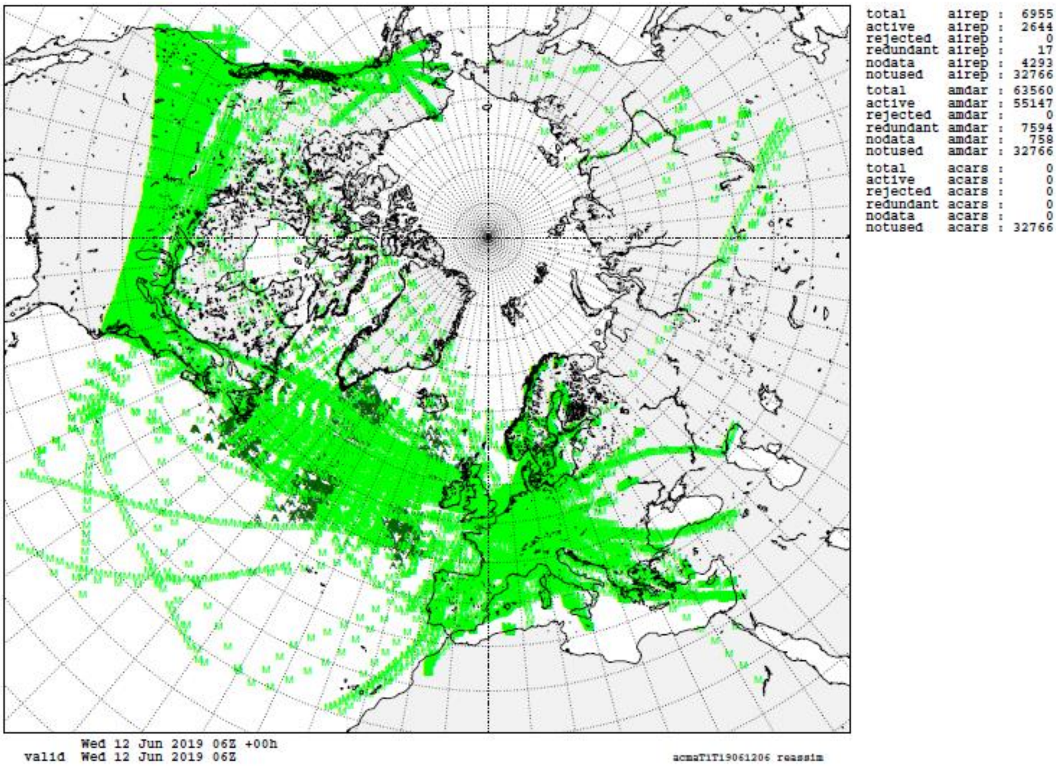

Figure 3.1.3 Upper air observations from aircrafts by June 12 2019

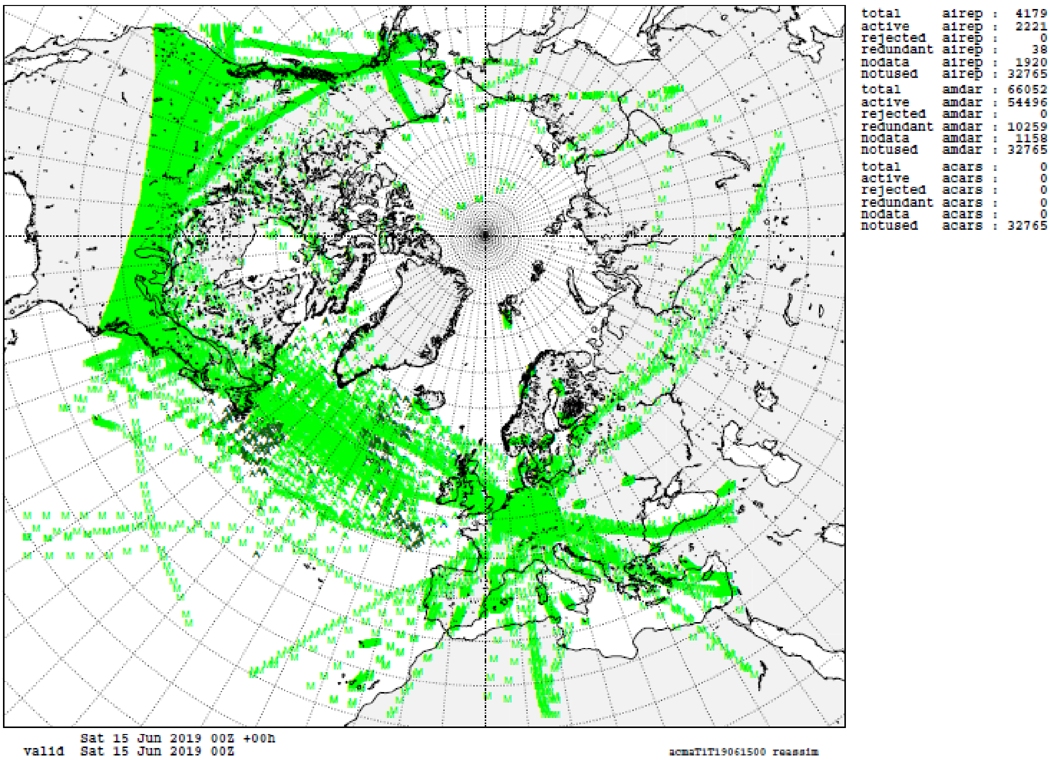

Figure 3.1.4 Upper air observations from aircrafts by June 15 2019

The GCOS Surface Network (GSN) and the GCOS Upper-Air Network (GUAN) were established in 1995 to serve the needs of global climate applications. The GSN network consists of 1023 stations while the GUAN network holds 177 stations.

These networks are a minimum configuration for global modelling purposes. The GUAN stations have a separation of 5-10 degrees latitude which should be sufficient to resolve synoptic scale waves. For other purposes, regional networks will have to be used as well. GUAN stations measure upper-air temperature, pressure (geopotential height) and wind as well as tropospheric humidity.

ECMWF provides monthly plots in order to monitor the actual availability of data from the GUAN stations (Figure 3.1.5); data with green squares fulfil the minimum requirements. Again, the spatial distribution of stations is low, especially for higher vertical levels. Upper-air information is important for data assimilation as well as for improving our understanding of basic atmospheric processes.

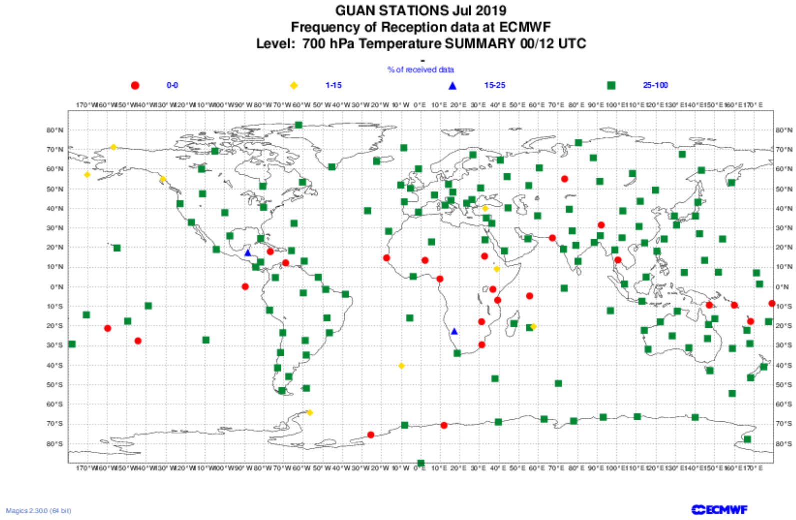

Figure 3.1.5 GUAN stations by July 2019: Frequency of data reception at ECMWF (700 hPa)

The International Arctic Buoy Programme (IABP) maintains a network of drifting buoys to provide meteorological and oceanographic data for real-time operational requirements and research purposes including support to the World Climate Research Programme (WCRP) and the World Weather Watch (WWW) Programme. An average of 25 buoys are in service at any time. The IABP drifting buoy data products described here are 12-hour interpolated pressure, temperature, position, and ice velocity grids available by year from 1979 through the present. Monthly ice buoy measurements in the Arctic may be found at https://iabp.apl.washington.edu/maps_monthly_map.html and are shown below. The buoy data are available from GTS and may also be downloaded directly from the webpage. When possible, the meteorological buoys are collocated with buoys measuring temperature, salinity and currents of the upper ocean.

The C3S project 311a Lot2 has been given the task to provide land and marine surface based meteorological records to the Copernicus Climate Data Store (CDS). Their work includes getting access to the data, converting it to a Common Data Model (CDM), performing quality assurance and finally providing it to the CDS. The comprehensive collection of surface meteorological observations will be made available from the CDS

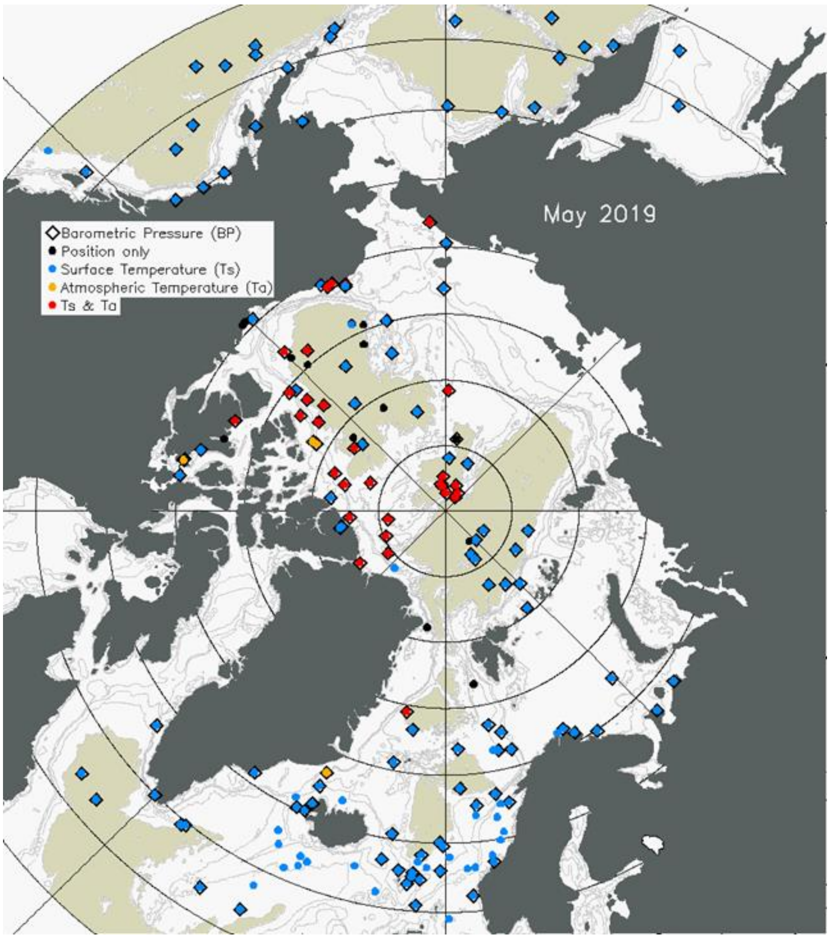

Figure 3.1.6 Monthly ice buoy data available from IABP for May 2019

The C3S_311a marine inventory includes all data in the International Comprehensive Ocean-atmosphere Data Set (ICOADS, https://icoads.noaa.gov/) which is a global ocean marine meteorological and surface ocean dataset that contains meteorological and oceanographic variables, such as sea surface and air temperatures, wind, pressure, humidity, and cloudiness. The ICOADS is based on a large number of national and international data sources and contains measurements and observations from ships (merchant, navy, research), moored and drifting buoys, coastal stations, and other marine and near-surface ocean platforms.

Whereas the marine meteorological observations have been collected and indexed (C3S_311a_Lot2 annex), the land observation data bases are less coordinated, and the C3S_311a inventory work on land meteorological observations is still ongoing. The Arctic stations currently in the inventory are shown in Figure 3.1.7 for daily stations. Most of these stations have been prioritized and will be part of the first full data release scheduled for September 2019, though slightly delayed. The first release will among other things include the Global Historical Climatology Network – Daily (GHCNd) data which are land based station data available from The National Centres for Environmental Information (NOAA) (www.ncdc.noaa.gov).

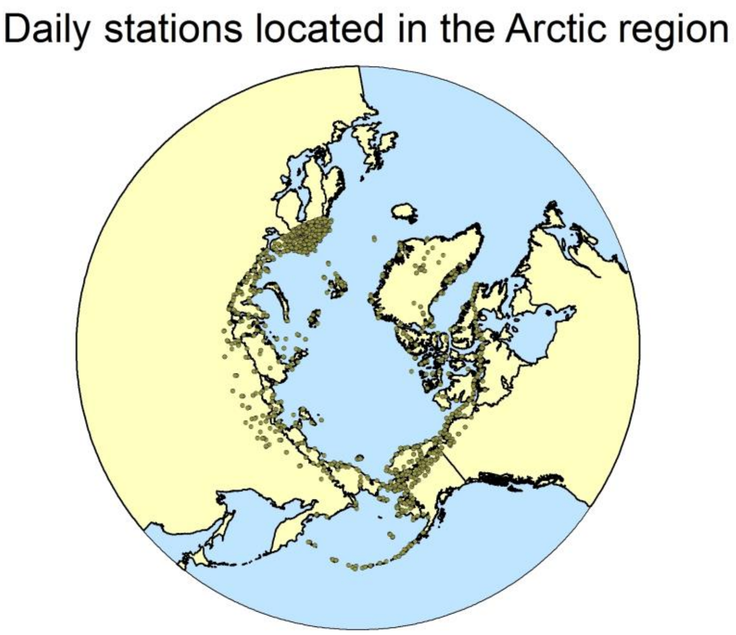

Figure 3.1.7: Daily stations for the CDS inventory located in the Arctic regions

The CDS will have daily updates for all operational sub-daily and daily stations. These will be appended to the station data with a caveat that no QC has been applied and that they are provisional. A complete refresh will be run periodically (Simon Noone, personal communication).

New data will be prioritized for processing through to the CDS based on variable availability, temporal coverage, location and metadata. The variables that will be prioritized in the C3S land inventory are shown in the table 3.1.1.

Table 3.1.1 Prioritized variables

Sub-daily	Daily	Monthly
Temperature Precipitation Mean sea level pressure Water vapor measurements Wind measurements Snow measurements	Temperature Precipitation Sunshine hours Mean sea level pressure Wind measurements Water vapor measurements Snow measurements	Temperature Precipitation Sunshine hours Mean sea level pressure Water vapor measurements

Some of the inventoried data sets contain variables not in the prioritized list (e.g. hail size, visibility, cloud ceiling). In this case, summary information about data availability (temporal and spatial coverage) is provided on a one-page information sheet to support more specialized applications.

All data and products in CDS are freely available. They are described using the ISO19115 metadata record standard and are made available through the protocols for interoperability with the World Meteorological Organization Information System (WIS) and the EU's INSPIRE initiative, respectively. All data and products will be assigned a Digital Object Identifier (DOI) (ECMWF Newsletter No. 151, Spring 2017).

The CDS will include in situ datasets for stations owned by national meteorological services and other operators who provide the data free of charge. As some operators do not have a full free and open data policy, a data sharing agreement between EUMETNET and Copernicus has been put in place, stating that “Copernicus Services can reuse all the data produced by EUMETNET members for their own purposes free of charge. The only limitations are attribution of ownership and proper licensing with data owners in case of redistribution of such data”. This means that data from WMO/GTS and similar operational delivery mechanisms can be provided to Copernicus services free of charge.

It will certainly improve the access to meteorological observation data, when all the collected data are made available from the CDS.

Baseline Surface Radiation Network (BSRN)

The BSRN is a world-wide network of radiation measurements with the highest achievable standards. At present 52 stations exist (see Figure 3.1.7), and six of these are situated in the Arctic (Lincoln Sea, Alaska, Russia, Siberia, Spitsbergen and the Shetland Islands). The sites of the stations have been selected to be as representative of the surrounding region as possible.

The data are redistributed via ftp and the web-based Data Publisher for Earth and Environmental Science, PANGAEA. Online data access is possible upon acceptance of data release guidelines.

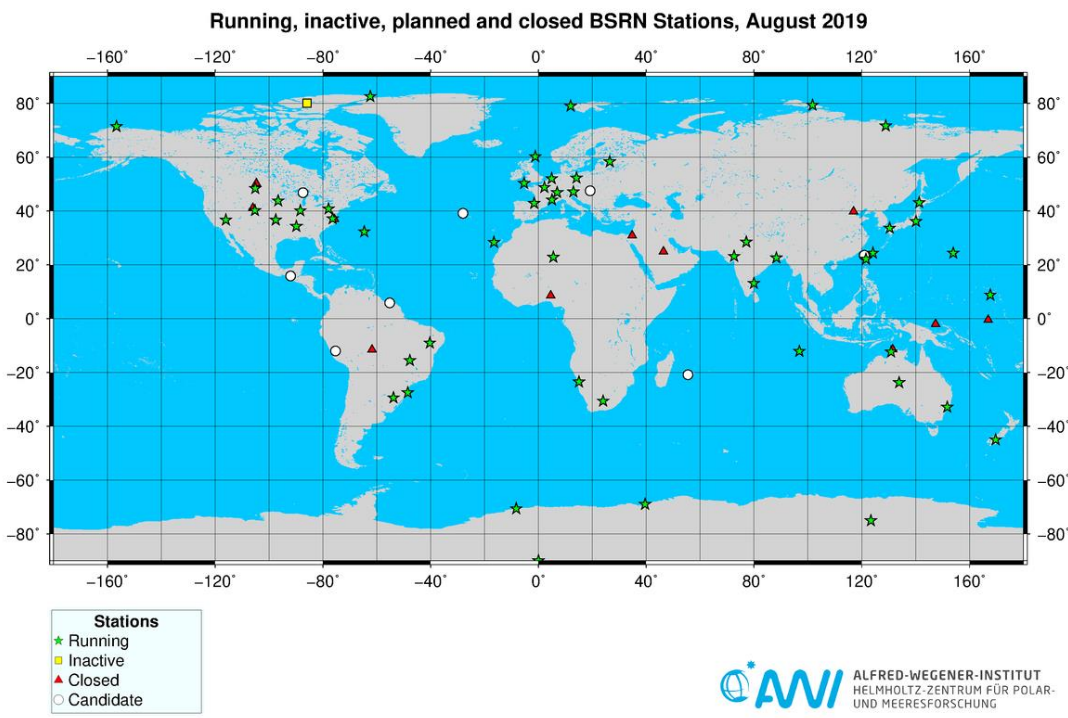

Figure 3.1.7 Running (green star), inactive (yellow square) and red triangle (closed) BSRN stations by December 2018.

Greenland data

The Programme for Monitoring of the Greenland Ice Sheet (PROMICE) (www.promice.dk) is operated by GEUS (Geological Survey of Denmark and Greenland). 25 Automatic Weather Stations exist in

Greenland, some of these supported by additional projects (e.g. the Greenland Analogue Project (GAP) and the Greenland Climate Research Centre (GCRC)). Delivery of data is not fully stable.

The Polar Portal (www.polarportal.dk) is a joint initiative of Danish institutions conducting research in the Arctic. These institutions are DMI, GEUS, DTU Space and DTU Byg. Data shown is from surface observations, satellites and climate models. Surface observations incorporated include the PROMICE stations on the ice sheet as well as DMI's stations along the coasts and on Summit. In addition, permafrost observations at selected locations are available. Numerous satellite products include the position of glacier fronts, mass and height budgets of the ice sheet, albedo, sea ice extent and the position of icebergs.

The Greenland Climate Network (GC-Net) consists of 18 Automated Weather Stations (AWS) distributed over the Greenland ice sheet. The GC-Net provide NRT meteorological and glaciological parameters at various locations on the ice sheet. The meteorological variables include air temperature, wind speed, humidity and pressure. Surface energy and mass balance are measured as well. Currently, the GC-NET archive contains more than 50 station years of measurements. These data have been quality controlled and calibrated.

3.1.2. Data with restricted availability

A large number of meteorological observations are performed as part of a national or community network and are only freely available upon request. Other data are collected as part of research projects in which resources are not available to make the data freely available.

Russian data

The Arctic and Antarctic Research Institute (AARI) in St. Petersburg (www.aari.ru) have weather observations from Tiksi (WMO station 21824), covering the period 1932 to near present and Ice Base Cape Baranova station (WMO station 20094), which is situated at 79°N, 10°E on Severnaya Zemlya Island. Data coverage is 2013 to near present.

The World Data Centre in Obninsk (www.meteo.ru) apparently has a large number of meteorological data that are available upon request. Actually, retrieving the data is less straightforward and requires some effort.

Greenland data

ASIAQ (Greenland Survey) runs a network of automatic weather stations, mostly located at airports, but also in some mining locations and a few research sites. Data may be obtained by request and are for instance used by the C3S Arctic Regional Reanalyses. Most of these stations are however close to SYNOP stations, and the delivery of data is not fully stable. It was noted that there is an opportunity to leverage ASIAQ's annual summer maintenance trips for sharing logistics.

Real-time data may also be obtained from airports. Their response to faults is however very slow according to experience by the C3S Arctic Reanalysis project.

3.2. Atmospheric composition

CAMS obtain data through the WMO GTS, from direct upload/download to/from FTP servers by collaborating organisations/projects/infrastructures (e.g. NILU, IAGOS, ACTRIS) and by download from databases on the web (e.g. AERONET, NDACC, WOUDC, WDCGG). The data formats include BUFR, NASA Ames, CSV, NetCDF and many others. The timeliness also varies from minutes to years. CAMS mostly rely on the observation networks to do the quality control. In order to support that process, CAMS has dedicated contracts (or is setting these up) with some of the major in situ networks and infrastructures (Integrated Carbon Observation System - ICOS, Network for the Detection of Atmospheric Composition Change - NDACC, In-Service Aircraft for a Global Observing System - IAGOS, EEA, European Aeroallergen Network - EAN, WMO-GAW, The Aerosols, Clouds and Trace gases Research Infrastructure – ACTRIS/EMEP). CAMS has also established some agreements with non-European partners, most notably with Environment Canada. Through the CAMS_84 contract on Evaluation and Quality Control there is some effort to make the processing of the various data formats more efficient.

3.2.1. Freely available data

This section summarises the atmospheric observations that are freely available to use. Most of these data are collected and used in CAMS, but additional datasets are listed as well.

The number of Arctic stations providing data for atmospheric research is low, but the measurements carried out at the various sites are relatively complex and cover a wide range of measurement principles and compounds, as seen from table 3.2.1. Public data are available from 11 stations, and ranging from point measurements on the ground (ground based in-situ) to remote sensing instruments, contributing to 12 national, regional or global monitoring networks.

Measurement principle	Compound	Stations	Network
Ground based in-situ, aerosol composition	Organic and inorganic, heavy metals	Alert, Nuuk, Summit, Zeppelin	AMAP, EMEP
Ground based in-situ, aerosol optical properties	Light absorption, light scatter and backscatter, vertical distribution, black carbon	Alert, Summit, Zeppelin	GAW, NOAA, ACTRIS
Ground based in-situ, aerosol physical properties	Number concentration, Size distribution	Alert, Zeppelin	GAW, NOAA, ACTRIS
Ground based in-situ gas phase	OVOCs, Halocarbons	Alert, Zeppelin	AMAP, GAW, EMEP, ACTRIS
Ground based in-situ	Ozone	Eureka, Alert, Nuuk, Summit, Zeppelin	GAW, NOAA, AMAP, EMEP
	Aerosol chemical species	Alert	AMAP
	OVOCs	Alert, Zeppelin	AMAP, GAW, EMEP
	NO2	Zeppelin	EMEP, AMAP, GAW
	CCNC	Zeppelin	GAW
	Size distribution	Zeppelin	ACTRIS, EMEP, GAW

	N_2O	Zeppelin	EMEP
	CH_4	Zeppelin	EMEP, ICOS
	CO_2	Zeppelin	GAW, ICOS
	CO	Zeppelin	EMEP, ICOS
	Halocarbons	Zeppelin	GAW, ACTRIS
Sonde	O_3	Alert, Eureka, Ny-Ålesund, Thule, Resolute, Goose Bay	NDACC, WOUDC
	Aerosol backscatter	Alert, Thule	NDACC
FTIR spectrometer, column and profile	C_2H_2, C_2H_6, CH_3OH, CH_4, $ClONO_2$, CO, HCl, HCN, HCOOH, HF, HNO_3, N_2O, O_3	Eureka, Ny-Ålesund, Thule	NDACC, WOUDC
LIDAR	Aerosol	Eureka, Ny-Ålesund, Thule	NDACC
	O_3	Eureka, Ny-Ålesund	NDACC
DOAS UV-VIS spectrometer, column and profile	NO_2, Aerosol, H_2CO, BrO, O_3	Eureka, Ny-Ålesund, Sodankylä	NDACC, Pandonia
SAOZ UV-VIS spectrometer, column	O_3, NO_2	Ny-Ålesund, Thule	NDACC
Microwave radiometer	ClO, O_3, N_2O	Ny-Ålesund, Thule	NDACC
Aircraft	O_3, CO	Arctic flights	IAGOS
	AOD	A few Acrtic stations	AERONET
Ship, Polarstern	CH_4, CO_2	Arctic Ocean	ICOS
FTIR TCCON	CO_2, CH_4, N_2O, HF, CO, H_2O, HDO	Eureka, Ny-Ålesund, Sondakylä	TCCON

Table 3.2.1. Overview of measurement principles, components, stations and networks in Arctic areas, relevant for CAMS and included in the current report.

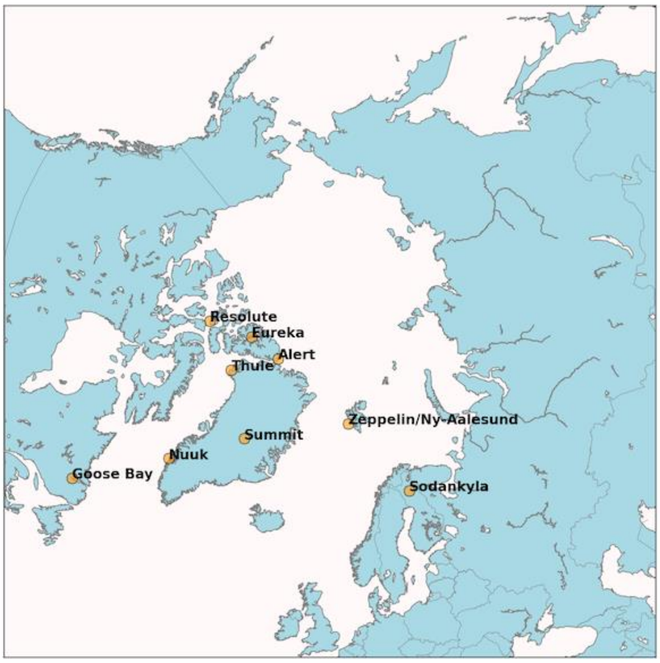

*Figure 3.2.1 The geographical distribution of atmospheric chemistry stations in the high-Arctic, providing free and open ground based and remote sensing in-situ data through various programs, networks and data archives (as listed in the 3rd column of table 3.2.*1)

ICOS is the Integrated Carbon Observation System, a European Research Infrastructure. ICOS has twelve member countries and operate over 130 greenhouse gas measuring stations aimed at quantifying and understanding the greenhouse gas balance of Europe and neighbouring regions. ICOS data is openly available at the Carbon Portal, a one-stop shop for all ICOS data products. https://www.icos-cp.eu/. For time being, ICOS Arctic data is available from two fixed observatories (Ny-Aalesund, Norway and Nord, Greenland) in addition to data from vessel Polarstern.

NDACC is the Network for the Detection for Stratospheric Change. NDACC is composed of more than 70 globally distributed, ground-based, remote-sensing research stations with more than 160 currently active instruments, providing high quality, consistent, standardized, long-term measurements of atmospheric temperatures and trace gases, particles, spectral UV radiation reaching the Earth's surface.

For NDACC, it is of value to offer data to the scientific community with a maximum delay of one month. If these rapid delivery data are of lower quality than traditional NDACC certified data, if the data have not yet been quality controlled, or if the data is less complete (e.g. missing uncertainty estimates) then these data must be identified as 'Rapid Delivery (RD)'. These are available separately on the NDACC public website. As soon as the standard verified version is available, the RD data will be removed and the fully verified version will be archived in the NDACC archive. Some RD datasets are from instruments that are not yet NDACC affiliated instruments. Data can be found at http://www.ndaccdemo.org/.

Figure 3.2.2 gives an overview of the Northern hemisphere high latitude NDACC stations. Within NDACC, there is, together with representatives from ESA, NASA, NILU and other scientific advisory groups, an ongoing effort to harmonise the data formats. CAMS supports NDACC through one of its in situ support contracts to increase the number of RD stations that can provide observations directly to CAMS.

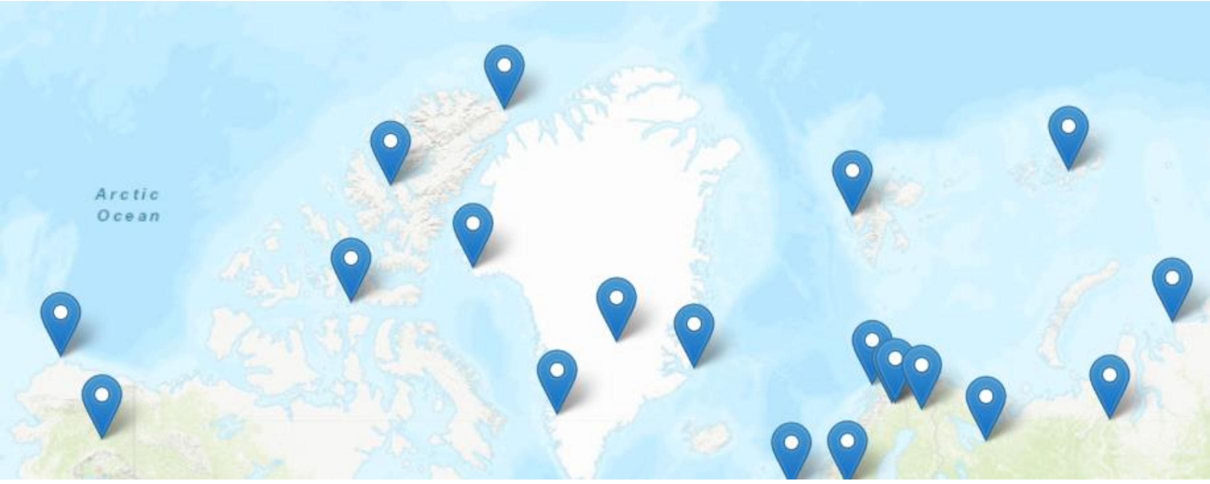

Figure 3.2.2 Overview of NDACC High latitude monitoring stations

IAGOS is the In-service Aircraft for a Global Observing System. IAGOS is a European Research Infrastructure for global observations of atmospheric composition from commercial aircraft. IAGOS combines the expertise of scientific institutions with the infrastructure of civil aviation in order to provide essential data on climate change and air quality at a global scale. In order to provide optimal information, two complementary systems have been implemented:

- IAGOS-CORE providing global coverage on a day-to-day basis of key observables
- IAGOS-CARIBIC providing a more in-depth and complex set of observations with lesser geographical and temporal coverage, https://www.iagos.org/iagos-data/

Use of the IAGOS data is free for non-commercial users. Access to the database is granted after registration and upon acceptance of the data protocol. The data include greenhouse gases (CO_2, CH_4, H^2O, N^2O, CFCs, SF6), reactive gases (O^3, CO, NOy, NOx, SO^2, VOC, HCOH), aerosol (number density, size distribution) and cloud particles (number density, size distribution). Data from flights covering the Arctic area is available, mainly from intercontinental flights crossing over Greenland.

The EBAS atmospheric database (http://ebas.nilu.no), originally designed for EMEP, archives contemporary data on atmospheric composition from ground stations around the globe as well as aircraft platforms. EBAS is used as the official data repository for the following frameworks:

- The Convention on Long-Range Transboundary Air Pollution - EMEP
- The WMO Global Atmosphere Watch Programme
- The Arctic Monitoring and Assessment Programme (AMAP)
- The EU-project Aerosols, Clouds, and Trace gases Research InfraStructure (ACTRIS)
- The OSPAR Convention – Comprehensive Atmospheric Monitoring Program
- HELCOM Commission

Most of the data stored in EBAS are originating from programs encouraging an unlimited and open data policy for non-commercial use. For scientific purposes, access to these data is unlimited and provided without charge. EBAS also operates a Near-Real-Time service for EMEP, GAW and ACTRIS.

The overview of EMEP, AMAP and ACTRIS stations in Arctic, as taken from EBAS, is found in Figure 3.2.3

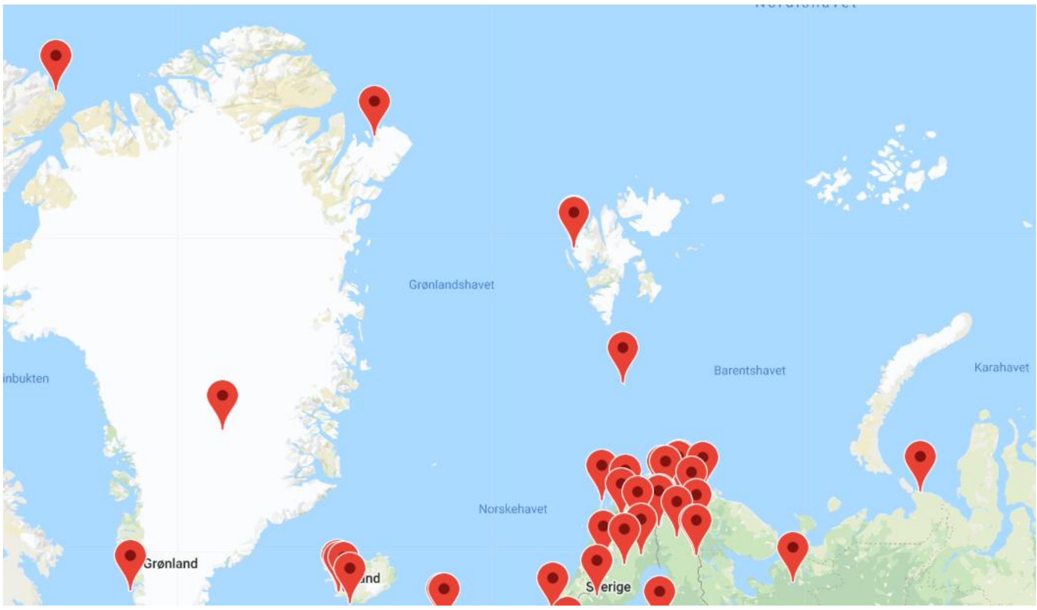

Figure 3.2.3. Overview of stations in Arctic that provides data to EMEP, AMAP and ACTRIS, stored in the EBAS data archive.

EMEP (http://emep.int) is the co-operative programme for monitoring and evaluation of the long-range transmission of air pollutants in Europe, a scientifically based and policy driven programme under the Convention on Long-range Transboundary Air Pollution (CLRTAP) for international co-operation to solve transboundary air pollution problems. The EMEP programme relies on three main elements:

- Collection of emission data,
- measurements of air and precipitation quality
- Modelling of atmospheric transport and deposition of air pollutions. The EMEP programme is carried out in collaboration with a broad network of scientists and national experts that contribute to the systematic collection, analysis and reporting of emission data, measurement data and integrated assessment results, where the co-ordination and intercalibration of chemical air quality and precipitation measurements are carried out at the Chemical Coordinating Centre (CCC).

AMAP is the Arctic Monitoring and Assessment Programme, and is one of six Working Groups of the Arctic Council. AMAP is mandated to:

- monitor and assess the status of the Arctic region with respect to pollution and climate change issues.
- document levels and trends, pathways and processes, and effects on ecosystems and humans, and propose actions to reduce associated threats for consideration by governments.
- produce sound science-based, policy-relevant assessments and public outreach products to inform policy and decision-making processes.

AMAP's work is directed by the Ministers of the Arctic Council and their Senior Arctic Officials, who have requested AMAP to also support international processes that work to reduce the global threats from contaminants and climate change. These include the UN Framework Convention on Climate Change, UNEP's Stockholm Convention on Persistent Organic Pollutants and Minamata Convention on mercury, and the United Nation's Economic Commission for Europe (UN ECE) Convention on Long-range Transboundary Air Pollution.

Since its establishment in 1991, AMAP has produced a series of high-quality reports and related communication products that detail the status of the Arctic with respect to climate and pollution issues and that include policy-relevant science-based advice to the Arctic Council and governments.

AMAP thematic data centres exist for:

- atmospheric contaminants data: at the Norwegian Institute for Air Research (NILU), Kjeller, Norway, and accessible through their EBAS database;
- marine contaminants data: at the International Council for the Exploration of the Sea (ICES), Copenhagen, Denmark, and accessible through their online EcoSystemData warehouse;
- radioactivity data, including both sources and levels and trends: at the Norwegian Radiation Protection Authority (NRPA), Oslo, Norway.

In addition, freshwater and terrestrial contaminants datasets have been compiled in the SynCon database. International data centres for ozone/UV, arctic ocean acidification, permafrost and other key AMAP-relevant data exist and their use is encouraged.

Other international data reporting initiatives are under development as part of the Sustaining Arctic Observing Networks (SAON) initiative.

ACTRIS is the European Research Infrastructure for the observation of Aerosol, Clouds and Trace Gases. ACTRIS is composed of observing stations, exploratory platforms, instrument calibration centres, and a data centre. ACTRIS serves a vast community of users working on atmospheric research, climate and Earth system and air quality models, satellite retrievals, weather analysis and forecast systems by offering high quality data and research infrastructure services for atmospheric aerosols, clouds, and trace gases.

The ACTRIS Data portal allows you to search, analyse, and download atmospheric composition data from a multitude of data archives. The data results from the activities of the ACTRIS infrastructure network complemented with data from other relevant networks, and gives free access to atmospheric observational data to analyse atmospheric composition.

Almost 135 different atmospheric variables are included in ACTRIS and about 65 sites are active. The measurements are done with 25 different methodologies with time resolution ranging from seconds to 1 week. The ACTRIS data management plan describes the data sets ACTRIS generates, how the data is made available, and the data repositories. The document also includes a list with all ACTRIS atmospheric variables together with recommended methodology. http://actris.nilu.no/ and http://ebas.nilu.no/

WMO-GAW is the World Meteorological Organization - Global Atmosphere Watch. WMO-GAW is addressing atmospheric composition on all scales: from global, regional, to local and urban. CAMS retrieved data from the GAW observational programs listed at https://public.wmo.int/en/programmes EBAS hosts the data of GAW – World Data Centre for Aerosols and GAW – World Data Centre for Reactive Gases. The GAW stations included in EBAS are to a large degree overlapping the EMEP and AMAP stations, but includes in addition the Alert and Barrow stations in North America, as seen from Figure 3.2.4

Figure 3.2.4 Overview of stations in Arctic that provides data to GAW-WDCA, stored in the EBAS data archive.

AERONET is the Aerosol Robotic Network. The AERONET project is a federation of ground-based remote sensing aerosol networks established by NASA and PHOTONS (PHOtométrie pour le Traitement

Opérationnel de Normalisation Satellitaire; Univ. of Lille 1, CNES, and CNRS-INSU) and is greatly enhanced by networks (e.g., RIMA, AeroSpan, AEROCAN, and CARSNET) and collaborators from national agencies, institutes, universities, individual scientists, and partners. For more than 25 years, the project has maintained a long-term, continuous and readily accessible public domain database of aerosol optical, microphysical and radiative properties for aerosol research and characterization, validation of satellite retrievals, and synergism with other databases. The network imposes standardization of instruments, calibration, processing and distribution. The data is publicly available and free of charge.

AERONET collaboration provides globally distributed observations of spectral aerosol optical depth (AOD). As seen from Figure 3.2.5 of the many AERONET sites around the world, only a very few are located in the Arctic.

WOUDC is the World Ozone and Ultraviolet Radiation Data Centre. WOUDC is one of six World Data Centres which are part of the Global Atmosphere Watch programme of the World Meteorological Organization. The WOUDC data centre is operated by the Meteorological Service of Canada, a branch of Environment and Climate Change Canada.

Ozone data set categories include total column ozone and vertical profile data from lidar measurements, ozonesonde flights, Ultraviolet (UV) radiation data set categories include broadband, multiband, and high-resolution spectral data types.

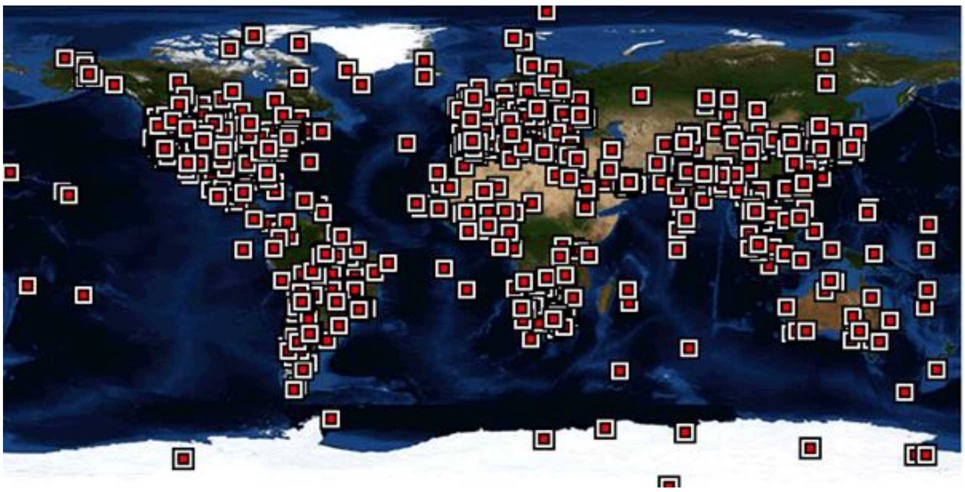

Figure 3.2.5 Overview of the AERONET sites

There are over 500 registered stations in the archive from over 150 contributors. Value added data products include total ozone time series graphs and near real-time ozone maps.

Compared to other parts of the world, the ozone measurement stations in the Arctic are rather sparse, however, the number of stations providing ozone sonde measurements is still large compared to stations providing data for other types of in-situ data, as seen from Figure 3.2.6

Access to the WOUDC data archive is available via numerous mechanisms in support of vendor neutral, multi-application access and integration. For scientific, educational and policy related use, access to WOUDC data is unlimited and provided without charge.

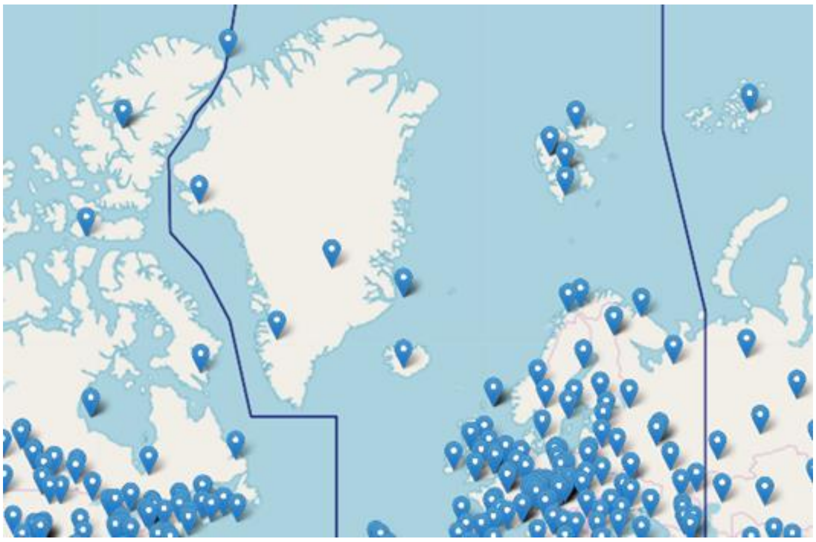

Figure 3.2.6 Overview of the WOUDC stations.

WDCGG is the World Data Centre for Greenhouse Gases. WDCGG is operated by the Japan Meteorological Agency (JMA) under the Global Atmosphere Watch (GAW) programme of the World Meteorological Organization (WMO).

Data Available at the Station(s) below:

CO_2	CH_4	N_2O	SF_6	SO_2F_2	NF_3	COS	$^{13}CH_4$
$^{13}CO_2$	$C^{18}O_2$	CH_3D	CCl_4	CH_3CCl_3	CFC-113	CFC-12	CFC-11
		CFC-115	CFC-114	CFC-13	HCFC-141b	HCFC-142b	HCFC-22
HCFC-124	HFC-125	HFC-23	HFC-152a	HFC-236fa	HFC-245fa	HFC-227ea	HFC-365mfc
HFC-134a	HFC-143a	HFC-32	HFC-4310mee	PFC-116	PFC-218	PFC-14	PFC-318
$CBrClF_2$	$CBrF_3$	$C_2Br_2F_4$	CO	$CHCl_3$	CH_3Cl	CH_2Cl_2	CH_3I
CH_3Br	CH_2Br_2	C_2HCl_3	C_2Cl_4	$CHBr_3$	H_2	$^{14}CO_2$	^{222}Rn
^{7}Be	TIC						

Figure 3.2.7 Parameters for in situ data provided through the WDCGG data centre.

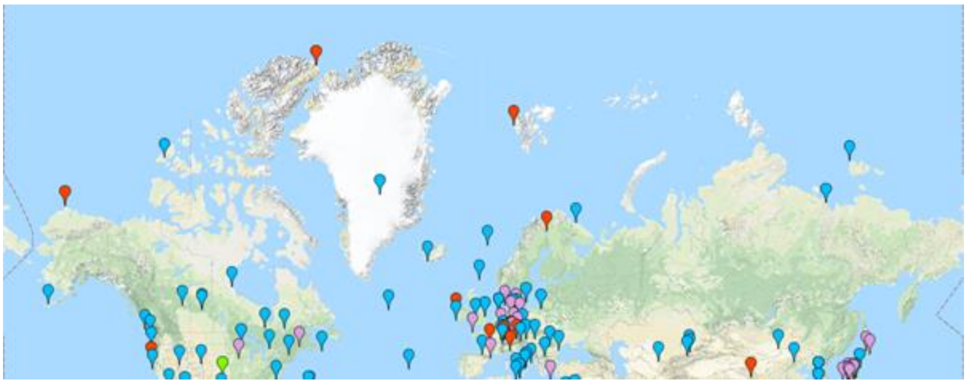

Figure 3.2.8 Overview of the WDCGG data and stations.

WDCGG collects, archives and distributes data provided by contributors on greenhouse gases (such as CO2, CH4, CFCs, N2O) and related gases (such as CO) in the atmosphere and elsewhere. All in situ variables available in WDCGG is listed in Figure 3.2.7, and the overview map of locations of the stations are found as Figure 3.2.8. As for the other data providers, the number of stations in the Arctic is rather sparse. Data is open, free of charge and can be downloaded from https://gaw.kishou.go.jp/search

IASOA is the International Arctic System for Observing the Atmosphere. IASOAs mission is to advance and coordinate research objectives from independent pan-Arctic atmospheric observatories. IASOA collaborates with Greenland, Finland, Norway, Russia, and Canada, see Figure 3.2.9 for overview of locations. The IASOA Data Portal targets continuous atmospheric data collected at an IASOA Observatory. It does not address data from campaigns, satellites or model output. The available atmospheric data is: air temperature, geopotential height, zonal wind, meridional wind and specific humidity.

CAMS-84 take use of surface ozone data from five different IASOA stations: Villum, Zeppelin, Tiksi, Eureka and Alert. The data are received from their primarily archives of the data providers.

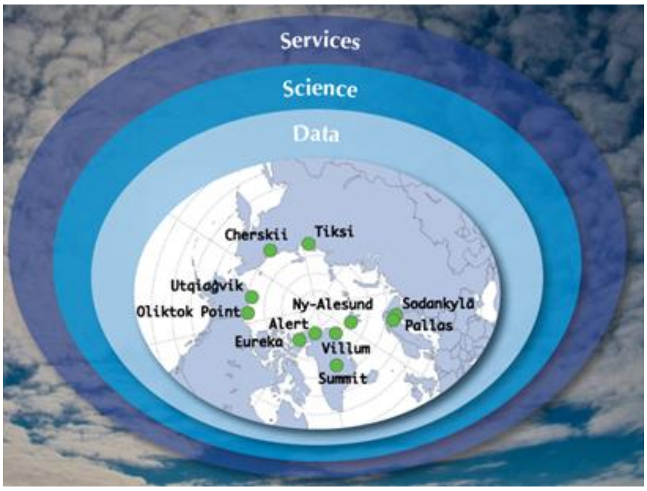

Figure 3.2.9 Overview of IASOA stations

EVDC – ESA **Atmospheric Validation Centre (EVDC)**; https://evdc.esa.int) provides O3 sonde data in NRT to CAMS. For the space component data primary collected for validation of satellite atmospheric composition products are available through the **EVDC.** The number of Arctic stations providing data to EVDC is around 10, all overlapping with NDACC and WOUDC.

Figure 3.2.10 Overview of the EVDC stations located in the Arctic

The EVDC holds a long list of atmospheric data, both historical and near-real time datasets, in the following categories, Table 3.2.2, that might be useful for CAMS, but are currently not used. Many datasets are overlapping with data summarised in Table 3.2.1 above.

Table 3.2.2 Categories of real time datasets in EVDC

Aircraft	asur, ch4tdl, fish, hagar, lidar.aerosol, lidar.watervapour, mas, mipas, safirea, uvvis.maxdoas
Assimilation	GOME
Balloon	atmoinspector, cryosampler, elhysa, far.ir.interferometer, fish, halox, lpma, salomon, sonde.backscatter, sonde.o3, sonde.ptu uvvis.saoz
Groundbased	ftir.c2h2, ftir.co2, ftir.hcooh, ftir.nh3, gps, pandora.no2, uvvis.doas.zenith.bro, ftir.c2h6, ftir.cof2 ftir.hf, ftir.no, lidar.o3, pandora.o3, uvvis.brewer,uvvis.doas.zenith.no2, ftir.ch3oh, ftir.h2co ftir.hno3, ftir.no2, lidar.temperature,radar.profiler, uvvis.doas.directsun.no2,uvvis.doas.zenith.o3, ftir.ch4,ftir.h2o,ftir.iso.h2o, ftir.o3, lidar.watervaporraman sodar, uvvis.doas.offaxis.aerosol , uvvis.dobson, ftir.clono2,ftir.hcl, ftir.iso.post.h2o, ftir.ocs, mwr.h2o, spectrometer, uvvis.doas.offaxis.h2co, uvvis.guv, ftir.co, ftir.hcn, ftir.n2o, ftir.tccon, mwr.o3, spectrophotometer, uvvis.doas.offaxis.no2, uvvis.saoz
Platform	filtration, hplc, photometer.cimel, photometer.cimel.seaprism, photometer.perkinelmer, radiometer.satlantic
Satellite	champ, cnofs, f3c.fm1, f3c.fm2, f3c.fm3, f3c.fm4, f3c.fm5, f3c.fm6, grace.a, poam3,sac.c
Ship	Filtration, hplc, photometer.perkinelmer, radiometer.satlantic, radiometer.trios, sea.atm.state, ftir.ch4, maeri, radiometer.biospherical, radiometer.simbada, satlanticsensor,spectroradiometer

Unexplored, non-quality assured data from the public and citizen's individual (private) instruments might be a future source of data. Data is currently openly available via internet portals such as e.g. Weather underground https://www.wunderground.com/ and NetAtMo. https://weathermap.netatmo.com/

3.2.2. Data with restricted availability

Observations that are not processed as near-real-time (NRT) often take a long time (more than a year) before they appear in a community data archive; this is in particular true for GAW observations.

The HALO database was established to hold and manage a wide range of data based on, or related to observations of the HALO research aircraft. The HALO database at DLR is available from https://halo-db.pa.op.dlr.de/

The primary data from uploaded datasets are available to all registered members of the particular mission. They may become publicly available after an embargo period.

Exchange of scientific data that are related to the HALO-community, other DLR research aircraft or in situ measurements: Long-term storage of data from a variety of research fields, e.g.

- Atmospheric chemistry and global pollution
- Atmospheric dynamics and transport
- Cloud research
- Meteorological research
- Climate research
- Global carbon cycle
- Polar Research
- Earth Observation
- Geophysics and Geodesy

This includes, for example, the WINDVAL aircraft campaigns and the MOSAIC campaigns are among the largest central Arctic campaigns ever, though there is limited information on data availability.

Surface ozone is measured by FMI at Pallas, and is used in CAMS-84. It would benefit the CAMS-84 service to have these data in near-real-time.

The Pandonia Global Network (PGN) is a collaborative effort between ESA and NASA with the aim of providing fiducial reference measurements for satellite validation and air quality monitoring. The main instrument of the PGN is the Pandora and Pandora-2S spectrometer system. The Pandora Spectrometer system is designed to specifically look at levels of ozone, nitrogen dioxide, formaldehyde and SO2 in the atmosphere.

PGN provides real-time standardized, calibrated and verified QA/QC air quality data. A Pandora-2S spectrometer was installed at the Sverdrup research station in Ny-Aalesund, Svalbard in August 2019. Located at 78N, this instrument will be the northernmost Pandora in the world, providing unique background data from the Arctic. https://www.pandonia-global-network.org/home/about/

3.3. Ocean

In situ marine observations in the Arctic are mainly made by Arctic countries (Canada, Denmark, Norway, Russia and USA) for national waters and by all countries with an interest in the Arctic and some international organizations for Arctic open waters. Depending on purpose and resource availability, some monitoring activities are regular, operational and sustainable (e.g. for purposes of operational services, environment and marine resource assessment and climate change) while the others are short term, irregular and less sustainable (e.g. for purposes of research and commercial interests).

The observations are collected by leading institutes/scientists of projects and programs, national, regional and global oceanographic data centres. A non-exhaustive list of the Arctic in-situ data collectors are given below:

National data centres

- Marine Environment Data, Fisheries and Oceans Canada, http://www.meds-sdmm.dfo-mpo.gc.ca/isdm-gdsi/index-eng.html
- ECDS - Environment Climate Data Sweden. https://ecds.se/
- Ifremer Coriolis. http://www.coriolis.eu.org/
- Russian Research institute of Hydrometeorological Information – World Data Centre, http://www.meteo.ru
- National Oceanographic Data Centre, USA. http//www.nodc.noaa.gov
- Norwegian Polar Data Centre at the Norwegian Polar Institute (https://data.npolar.no/home/
- Chinese National Arctic and Antarctic data centre. http://www.chinare.org.cn/index/
- Korean Polar Data Centre https://kpdc.kopri.re.kr/about/intro
- Arctic Data Archive System, Japan. https://ads.nipr.ac.jp/
- Arctic Office at British Antarctic Survey, UK. https://www.bas.ac.uk/data/our-data/data-systems/

Institutional databases

- Oceanographic Databases, Bedford Institute of Oceanography. http://www.bio.gc.ca/science/data-donnees/base/index-en.php
- ArcticNet: http://www.arcticnet.ulaval.ca/data-management
- The Greenland Ecosystem Monitoring (GEM) Database, https://data.g-e-m.dk/
- Department of Bioscience, Aarhus University, http://bios.au.dk/en/consultancy/scienticdatacentres/
- Data portal of the Alfred Wegener Institute (AWI) Helmholtz Centre. https://data.awi.de/?site=home
- UDASH – Unified Database for Arctic and Subarctic Hydrography, https://epic.awi.de/id/eprint/47869/
- Integrated Climate Data Center - ICDC. https://icdc.cen.uni-hamburg.de/1/daten/ocean.html

International data network or data portal

- International Arctic Science Committee (IASC) Arctic Data Committee. https://arcticdc.org/
- JCOMM operational observations (JCOMMOPS). http://www.jcommops.org/board
- Norwegian Marine Data Centre, http://metadata.nmdc.no/UserInterface/#/
- Arctic and Antarctic Research Institute and NIERSC St.-Petersburg, www.aari.ru
- Arctic ROOS. https://arctic-roos.org/
- SeaDataNet, https://www.seadatanet.org/
- ICES, http://www.ices.dk
- CMEMS INSTAC: http://marine.copernicus.eu/
- EMODnet-European Marine Observation Data Network, http://www.emodnet.eu
- IASC Network on Arctic Glaciology. https://nag.iasc.info/data
- The Global Sea Level Observing System (GLOSS). https://www.gloss-sealevel.org

- Permanent Service for Mean Sea Level (PSMSL). https://www.psmsl.org
- OceanSITES. http://www.oceansites.org/data/index.html
- The Arctic Monitoring and Assessment Programme (AMAP) - www.amap.no
- Conservation of Arctic Flora and Fauna (CAFF) - www.caff.is
- The Circumpolar Biodiversity Monitoring Program (CBMP) - www.caff.is/
- Protection of the Arctic Marine Environment (PAME) - www.pame.is
- The Sustaining Arctic Observing Networks (SAON) - www.arcticobserving.org
- Arctic Data – http://portal.intermap.com
- European Ocean Biogeographic Information System - www.eurobis.org

Research databases

- Web-based Data Publisher for Earth and Environmental Science, PANGAEA, https://www.pangaea.de/
- Arctic Portal. https://arcticportal.org/science/arctic-data
- NSF Arctic data centre. https://arcticdata.io/catalog/

The H2020 INTAROS project provided an assessment of current Arctic in-situ data availabilities, together with a gap analysis (INTAROS, 2018). Unfortunately, the assessment only covers in-situ data measured or collected by project partners, i.e., observations from Norway (NPI, IMR, NIVA, NORUT, UiB, UNIS, NERSC), Denmark (AU, DTU), Poland (IOPAN), France (EGO), Canada (SIO acoustic propagation experiment) and Germany (AWI, FRAM and UDASH). A large number of databases and data centres have not been included in the INTAROS assessment. The parameters evaluated are mainly water temperature, salinity, sea level, currents, acoustics, oxygen and chl-a. There are a few datasets which have phytoplankton and optical measurements but the are not sufficient to form a basis for evaluation.

The CMEMS In Situ TAC (INSTAC) is the component of the Copernicus Marine Service which ensures a consistent and reliable access to a range of in situ data for the purpose of service production and validation.

INSTAC has two main objectives:

- To collect multi-source, multi-platform, heterogenous data, perform consistent quality control, and to distribute it in a common format (NetCDF) in near-real-time (within 24 hours) to the CMEMS Marine Forecasting Centre's (MFC) for assimilation into their numerical ocean models.
- To supply the MFCs and downstream users with re-processed 25-50-year products in delayed mode. In addition to the near-real-time products, these delayed-mode products are useful for model validation or assimilation in ocean reanalysis and climate studies

INSTAC is a robust distributed system operated by professional centres that have been involved in in-situ data management activities for decades. The system has proven its reliability with a mean availability of the data better than 99% over the past 4 years both at global and regional scales. The establishment of a connection to national in situ data providers together with a close collaboration with EMODnet Physics has increased the number of available platforms at global scale, from 1000 per day in 2008 to around 3500 in 2018

3.3.1 Data with open access

Data with open access will be evaluated via variables in this section.

3.3.1.1 Water temperature and salinity

Profile measurements

T/S profiles in the Arctic can be measured by using bottle samples, CTD cast, XBT/BT/XCDT cast, Autonomous Pinniped Bathythermograph (APB), Argo floats, gliders, UAVs, towed CTDs and moorings. Our assessment will start from observations from well-known European and global databases, then extend to national and institutional databases and then research databases. In this assessment we will use the year 2016 as an example to evaluate the annual availability of data.

T/S profiles from SeaDataNet, ICES, EMODnet and WOD (World Ocean Database) 2018

SeaDataNet is an EU data initiative which collects historical observations from EU national oceanographic data centres. Metadata information can be obtained from http://SeaDataNet.maris2.nl/. Taking 2016 as an example, the total amount of datasets (i.e. number of profiles) is 4135. A geographical distribution of all datasets in the Arctic (Latitude > 60 N) is displayed in Figure 3.3.1. The data are mainly distributed in the Greenlandic and Nordic Seas. The data contributors are mainly from Norway, Germany, France, Iceland, UK, Denmark and Canada.

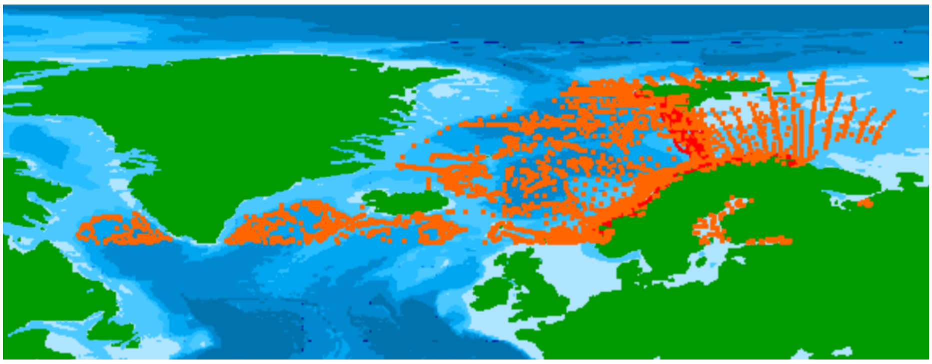

Figure 3.3.1 Geographic distribution of T/S profile monitoring stations in SeaDataNet during the period of Jan.-Dec. 2016

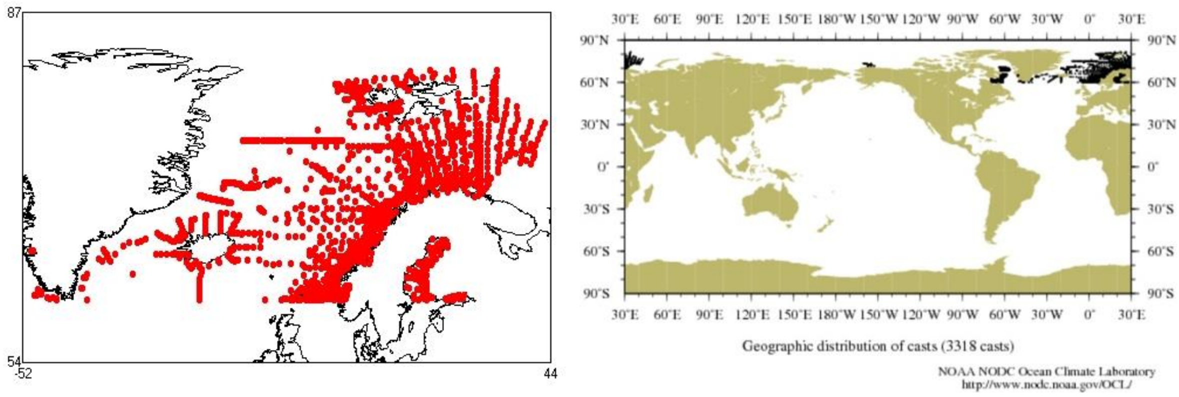

Figure 3.3.2 Geographic distribution of T/S profile monitoring stations from research vessels in ICES (left) and WOD 2018 (right) databases during the period of Jan.-Dec. 2016

The International Council for the Exploration of the Seas (ICES) collects data for fishery management in the North Atlantic. The data are disseminated through http://www.ices.dk . For the period of 2016, ICES observations cover mainly the Nordic Seas (52°W-48°E). The total amount of T/S profiles is 3153.

Most of the observations are to the east of 3°E. By comparing the data contributors and amount of contribution per country to ICES and SeaDataNet, it is found that the overlap between the two datasets is smaller than 25%. This leads to a conservative estimate of the total number of available T/S profiles in 2016 of 5466 from the ICES and SeaDataNet.

In EMODnet there are 3979 CTD profiles collected in 2016 in Arctic waters excluding Nordic Seas and Russian waters. The data are mainly collected from NMDC, WHOI, Sweden and PANGAEA

At the National Oceanographic Data Centre (NODC) in USA, a comprehensive World Ocean Database (WOD) can be found. As shown Figure 3.3.2 (right), there are 3318 profiles in WOD in year 2016 collected from research vessels. The station distribution of WOD is very similar to ICES in Nordic Seas. Hence one may assume that the additional 165 profiles are from US and Canadian coastal waters.

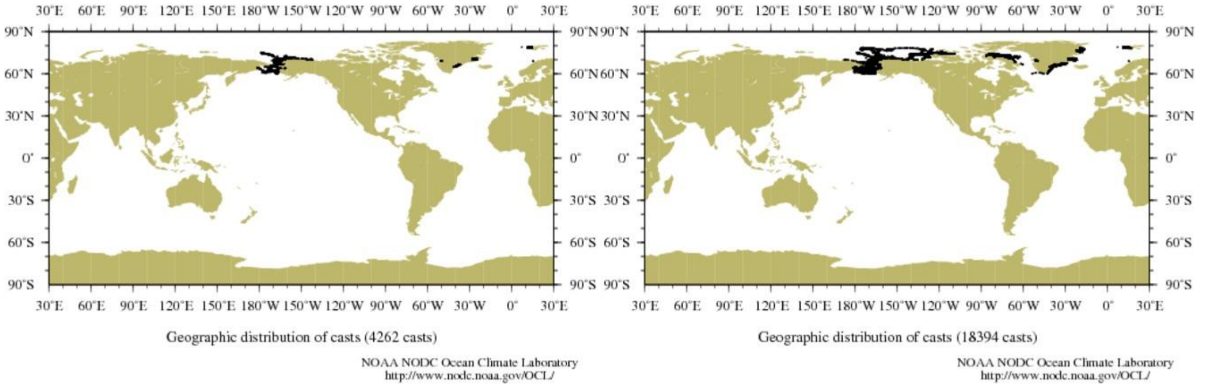

Figure 3.3.3 Geographic distribution of T/S profile monitoring stations from Autonomous Pinniped Bathythermograph *in WOD for 2016 (left) and 2016-2018 (right)*

In addition, WOD collects T/S profiles through Autonomous Pinniped Bathythermograph (APB) (Figure 3.3.3) and profile float (Figure 3.3.4), which have 4262 and 4136 profiles in 2016, respectively. The number of APB data has been increasing in recent years. In 2016-2018, the total number of APB profiles is 18394 which means that the APB data has increased by 66% from 2016 to 2017-2018 (6131 per annuum). Spatial coverage of the sensors, mainly located in the Marginal Ice Zone, is also significantly enlarged (Figure 3.3.3). The float profiles in the WOD database are mainly distributed in the Greenlandic and Nordic Seas (Figure 3.3.4)

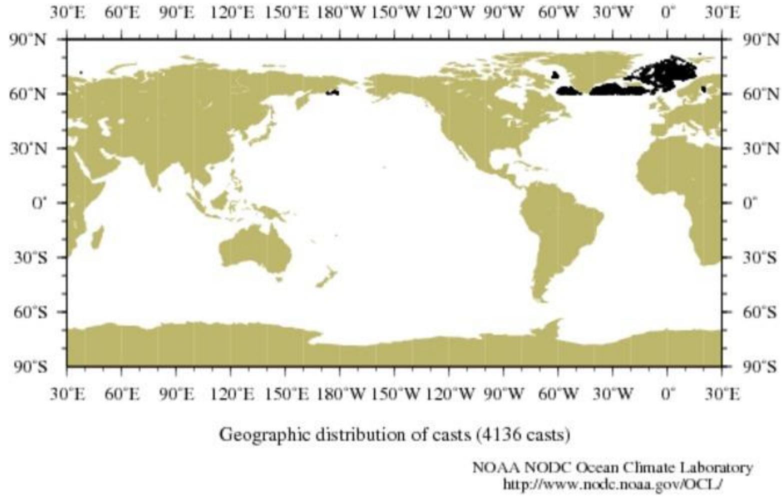

Figure 3.3.4 Geographic distribution of T/S profile monitoring stations from float profilers in WOD 2018 databases during the period of Jan.-Dec. 2016

Based on the above analysis, one can conclude that the number of T/S profiles per annum from SeaDataNet, ICES, EMODnet and WOD can reach 18008-19877. It should be noted that this does not include WOD data from gliders and drift buoys.

T/S profiles from moorings and ITPs

Moorings provide high frequency T/S observations. There are quite a few long-term research projects deploying moorings, for example

- Station Mike, Nordic Sea: 1 mooring
- North of Svalbard (UIB): minimum 1 mooring
- A-TWAIN mooring array north of Svalbard: 2-3 moorings (Norwegian Polar Data Centre)
- Barents Sea Opening Mooring Array: 5 moorings (IMR-NMDC)
- FRAM Program: 3-5 moorings (79 N, PANGAEA)
- A-TWAIN: Deep-ocean moorings north of Svalbard: 1-2 moorings
- Greenland Ecosystem Monitoring Programme: 1 mooring (AU)
- UNIS ocean observing system: varying, up to 5 moorings (UNIS)
- NABOS II (Nansen and Amundsen Basins Observational System): up to 13 moorings. (NSF) https://nabos.iarc.uaf.edu/NABOS2/data/registered/2018/moorings.php
- Churchill Marine Observatory - Environmental Observing CMO-EO: moorings in NW shipping route, Hudson Bay, >5 moorings
- ALTIMA Moorings in Chukchi and western Beaufort Seas: 7-12 moorings, (NSF)
- Deadhorse/Endicott, Beaufort Sea, 1 mooring (NSF)
- IERP moorings in Chukchi Sea: 3 moorings (NSF)
- OceanSITES north of 60 N: 31 (29 in the Nordic Seas and 2 in Bering Strait)

An introduction to most of the moorings in Nordic Seas can be found in INTAROS deliverable D2.1 (INTAROS, 2018). Metadata of moorings from NSF funded projects can be found in NSF Arctic Data Centre. OceanSITES website has a detailed description on their moorings.

A rough calculation gives an estimate of around 65 moorings operating in the Arctic every year. These platforms can generate significant amount of hourly profiles. They measure not only T/S profiles, but also other parameters, e.g. currents and/or biogeochemical parameters. However, most of them are offline, and it is difficult to maintain these moorings in central Arctic.

The ITPs can measure temperature or T/S profiles in the upper 800 meters. WHOI has been operating ITPs for more than 10 years. Currently there are still 5 ITPs operated by WHOI in the central Arctic. It is expected that there will be up to 17 ITPs deployed in the next few years.

T/S profiles from other sources

In addition, there are significant amount of data from national data centres in China, Japan, Korean, Canada and USA etc., and also from research projects, which have not been included in the above described databases (SeaDataNet, ICES, EMODnet, WOD, moorings and ITPs).

For example, China has carried out 9 independent Arctic expeditions and several joint expeditions since 1999. In recent years, the number of Arctic expeditions from China has been significantly increased.

For example, in 2018, China carried out 3 expeditions: the 9th Chinese Arctic Expedition, the 1st Sino-Russian joint expedition and the Nordic Sea expedition. The data from the expeditions are not immediately released; currently data from expeditions up to 2014 is available.

US NSF also operates an extensive Arctic research database – NSF Arctic Data Centre. Data from most of the NSF and NOAA funded projects in the Arctic can be accessed here.

For historical T/S profiles, in addition to SeaDataNet, ICES, EMODnet and WOD database, UDASH database also provide comprehensive data. It is a unified and high-quality temperature and salinity data set for the Arctic Ocean and the subpolar seas north of 65 °N for the period 1980-2015, generated by AWI. The archive aims at including all publicly available data and so far, consists of 288.532 oceanographic profiles measured mainly with conductivity/temperature/depth (CTD) probes, bottles, mechanical thermographs and expendable thermographs (Figure 3.3.5). To achieve a uniform quality level, approximately 74 million single measurements of temperature and salinity were thoroughly quality-checked. Data outliers, suspicious gradients and other suspect data were flagged for quick identification. The final archive provides a simple way of accessing most of the available temperature and salinity data for the Arctic Mediterranean Sea and represents a unique tool for a wide range of oceanographic analyses. More details of UDASH data can be found in Behrendt et al. (2018).

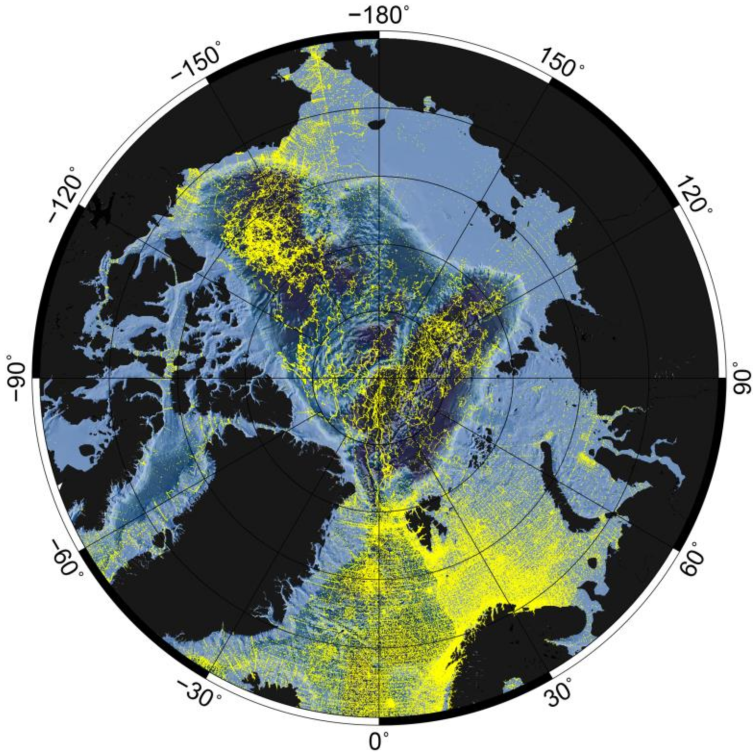

Figure 3.3.5 UDASH data coverage after quality check and data validation. Size: 288 532 profiles. Period: 1980–2015. Region: north of 65 °N.

SST and Sea Surface Salinities (SSS)

SST and SSS are measured not only by T/S profile monitoring platforms but also by surface drifters, moorings and Ferrybox systems.

Surface drifters: A drifter is composed of a surface float, which includes a transmitter to relay data via satellite, and a thermometer that reads temperature a few centimetres below the air-sea interface.

The surface float is tethered to a holey sock drogue (a.k.a. “sea anchor”), cantered at 15 m depth. The drifter follows the ocean surface current flow integrated over the drogue depth. It provides near real time (NRT), high quality SST measurements which are used in CMEMS for validating NRT forecast. Approximately half of the drifters also measure air pressure, and send the data to weather centres for improved marine forecasts. A smaller number of drifters have been deployed to measure other properties such as surface salinity and heat content in the upper 150m of the ocean. A snapshot of available surface drifters is shown in Figure 3.3.6.

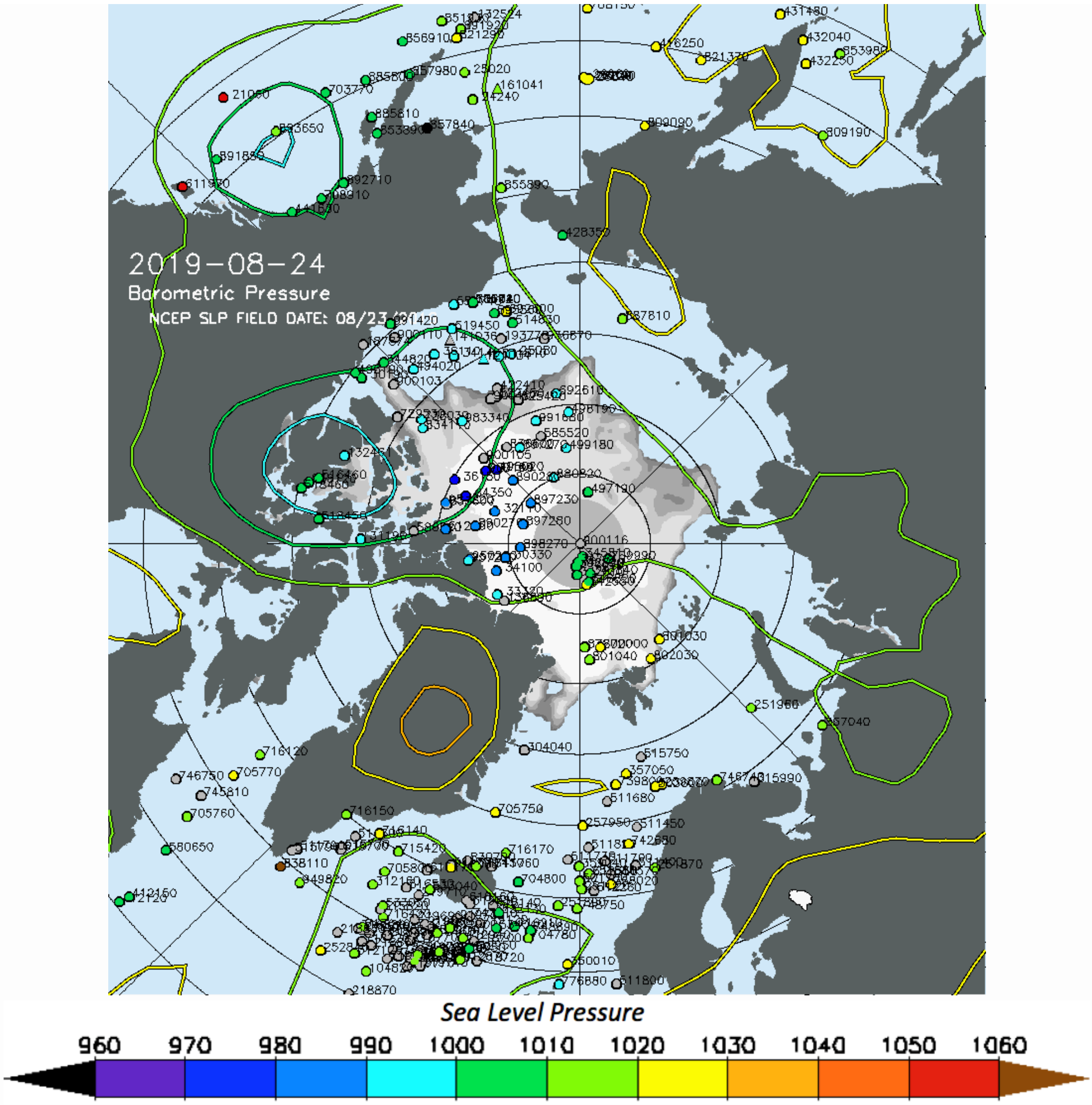

Figure 3.3.6 Surface drifter coverage on 24th August 2019, as reported in International Arctic Buoy Program (IABP). The contours show sea level pressure in colors.

In total there are 176 buoys in operation in the north of 60°N. In addition to classic SVP drifters and BD2GHI drifters, IABP also includes surface temperature, air temperature and sea level pressure reports from 5 ITPs, 3 IMBs, 3 UPTempOs, 8 snow buoys, 32 different kinds of ice buoys, 2 saildrones, 2 sidekick buoy, 2 SPIT buoy and 1 WARM buoy. These buoys may measure additional parameters but data are managed via different entities.

NDBC (National Data Buoy Centre) buoys: there are a few moored weather buoys measuring SST and reporting to NDMC database (https://www.ndbc.noaa.gov/).

Ferrybox lines operated in the Arctic: Ferrybox is an operational method to continuously measure physical, chemical and biological parameters on board of ships-of-opportunity. Ferrybox is a through-flow system installed on board of a ship-of-opportunity (ferry, cargo ship or research vessel) to measure automatically, continuously and unattended a series of important marine parameters. The sensors installed offer the opportunity to measure physical (salinity, temperature, turbidity), chemical (nutrients, pH, O_2, CO_2, DOC, oxygen) and biological (chlorophyll, phytoplankton composition, dominant/harmful species) parameters. Currently there are 5 Ferrybox lines operated in the Nordic Seas, Figure 3.3.7. Most of the Ferrybox lines measure temperature, salinity, O_2, chl-a fluoresces and turbidity at 4-5m depth. The amount of data, due to spatial continuous sampling, is significant when comparing with Argo and ship data. In principle on the effective sampling density of a Ferrybox is similar to a surface sampling from a mooring buoy. The Ferrybox data have already been collected in CMEMS INSTAC but rarely been used in CMEMS applications so far.

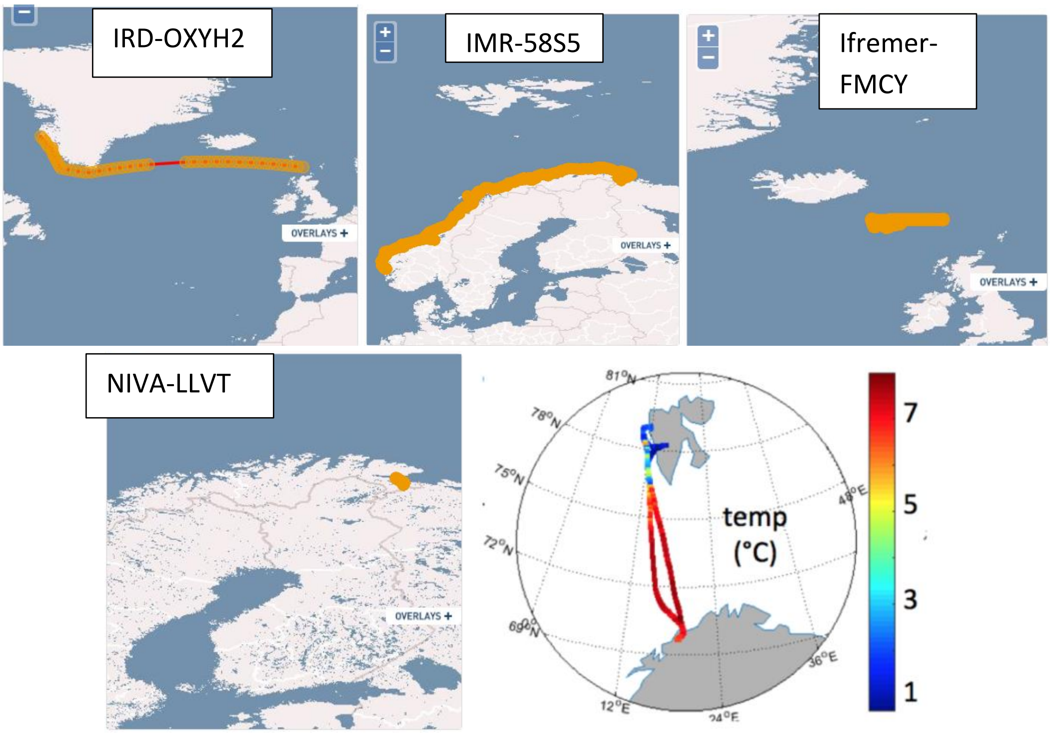

Figure 3.3.7 Ferrybox operated in the Nordic Seas

3.3.1.2 Currents

Current profiles: there have been many research projects deploying moorings and ship-board ADCPs to measure current profiles. Many of the on-going mooring projects (as listed in s3.3.1.1) such as FRAM, NABOS, ALTIMA, IMR moorings in the Barents Sea Opening and OceanSITES moorings. Research vessels are often quipped with Vessel-Mounted ADCP (VMADCP), e.g. in FRAM and AREX (Long-term large-scale multidisciplinary Arctic monitoring program). For the historical period, there is also a collection of following data:

- Nine mooring sites in northern Baffin Bay, operating in 1997-1998
- ArcticNet: Thirty-six moorings were deployed in the Beaufort Sea between 2004 and 2010. Mooring sites included the Mackenzie Shelf (North-Western to North-Eastern), Mackenzie Shelf slope, Kugmallit Valley, Canada Basin, Mackenzie Canyon, Amundsen Gulf (West to East), and outer Franklin Bay area. Bottom depth at deployment sites varied between 200 and 545 m. Mooring lines were equipped with various oceanographic instruments attached at different depths from 12 m to a maximum of 500 m below the surface. Moored instruments included ALEC conductivity-temperature-turbidity- chlorophyll-PAR sensors, SeaBird SB 26 and SB37 conductivity-temperature sensors, RMC4, RMC7 and RCM11 current meters, Nortek current profilers, McLane moored profilers and NIPR and Technicap sediment traps.
- Ocean current and sea ice statistics for Davis Strait: 2004-2010: ADCP and ice draft, 9 stations
- Two moorings BA01-05 and BA01-06 in Baffin Bay, operating in 2005-2008; 2006-2008

National Arctic expedition often measure currents. For example, in Japan Arctic Data Centre ADS, there are following entries which may contain currents observations:

- Acoustic Doppler Current Profiler (ADCP) data in Oshoro-Maru C40 cruise. [A20190805-001] ADS
- Acoustic Doppler Current Profiler (ADCP) data in Oshoro-Maru C56 cruise [A20190805-002] ADS
- Ocean current measurement in fjord in front of Bowdoin Glacier, Greenland in August 2014 [A20140829-002] ADS
- ESC12 mooring on the East Siberian Sea shelf (2012 deployment -2013 recovery) from IBRV Araon. [A20140425-002] ADS
- ESS12 mooring on the East Siberian Sea slope (2012 deployment -2013 recovery) from IBRV Araon. [A20140425-001] ADS
- GRENE Arctic Moorings (GAM) in the Northwind and Chukchi Abyssal Plain (2012 deployment -2013 recovery) [A20140425-003] ADS
- Mooring deployment at Bowdoin Fjord, Greenland in July 2016 [A20170420-003] ADS
- Mooring deployment at Bowdoin Fjord, Greenland in July 2017 [A20190517-008] ADS
- Mooring deployment at Bowdoin Fjord, Greenland in July 2018 [A20181001-001] ADS
- ROV observation in northern Bering Sea and Chukchi Sea. [A20131009-007] ADS

in 2007-2010 measuring velocity (both surface and profiles), temperature, salinity, and the vertical turbulent fluxes of heat, salt, and momentum in the ocean surface mixed layer with a custom-built sensor package about 5 m below the ice-ocean interface. In total 18 buoys were deployed.

Spatial distribution of these two categories of data entries are shown in Figure 3.3.8.

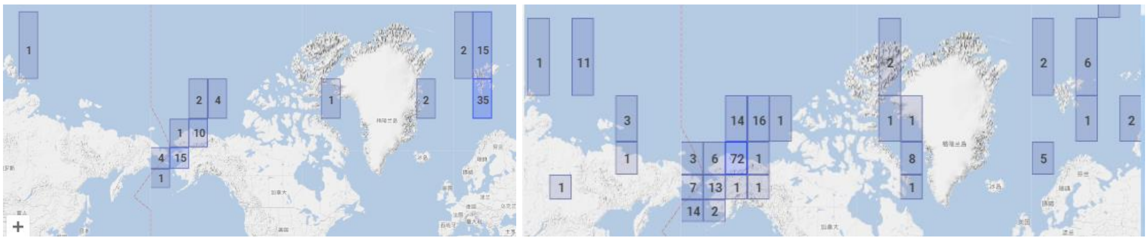

Figure 3.3.8 Spatial distribution of number of datasets containing key word "current meter" (left) and "ADCP" (right).

Surface currents: In addition to current meters and ADCPs, surface currents can also be derived from drift buoy data (as described above) and even Argo floats.

3.3.1.3 Sea level and waves

Sea level:

NRT data: Global Sea Level Observing System (GLOSS) collects and processes in situ sea level data from the Arctic Region. GLOSS has data from 10 tide gauge stations on the Arctic coast, providing sea level data with a frequency of 5-60 minutes. Original sea level data are measured at tide gauge stations from Arctic countries: Canada (>10), Denmark (4), Norway (17), Russian (13) and USA (10). In total there are more than 54 tide gauges (Tab. 3.3.1b) providing NRT sea level measurements.

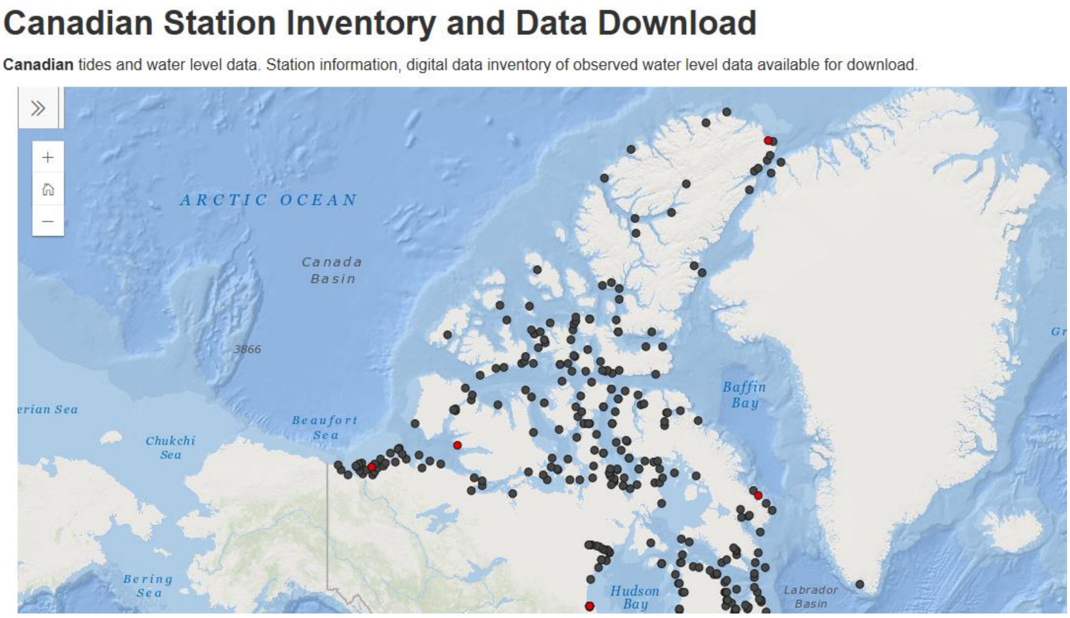

Figure 3.3.9 Tide gauge station distribution in Canadian tide and water level data archive.

Historical data: Canadian MED (Marine Environment Data Section) has a tide and water level data archive, currently containing sea level data from 188 stations in Canadian coast north of 60 N, with the earliest dating back to 1848. The number of observations spans on average 6 years per station, with 60 stations measuring water levels for over 50 years (Figure 3.3.9)

PSMSL has monthly mean sea level data from 42 Arctic tide gauge stations, which have been pre-processed to common local reference level.

Waves

Due to the presence of ice, it is a challenge to make wave observations in the Arctic water. However, due to accelerated melting of Arctic ice, waves are playing a more and more important role in the Arctic.

<u>*NRT data*</u>:

- NDBC database: currently there are 6 wave buoys reporting data in NDBC (https://www.ndbc.noaa.gov/) in the north of 60 N - one in Icelandic offshore and the other 5 in the open water area NW of Scotland.

- Canadian data: there are also two wave buoys operating in the Great Slave Lake. In Norwegian offshore there exist about 26 wave buoys operated by oil companies. The data access is under condition of approval from the owners.

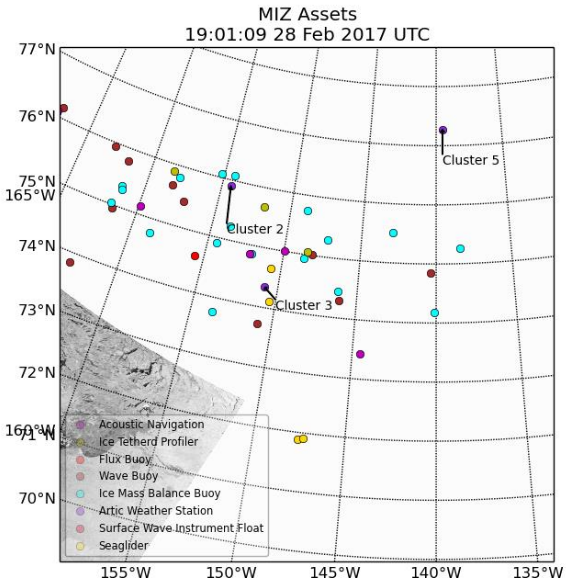

Figure 3.3.10 Station distribution of MIZ project

Historical data:

- Data from ONR (Office of Naval Research) Marginal Ice Zone (MIZ) Program: the purpose of MIZ is to investigate ice-ocean-wave-atmosphere interaction in the Beaufort Sea with models and a network of autonomous systems including wave buoys. The programme included 25 Wave Buoys to quantify open ocean and in-ice wave characteristics and evolution. 20 buoys were deployed in the summer period, 5 in the winter to measure directional wave spectra in the ocean. A station map is given in Figure 3.3.10. In addition, a network of ITPs, IMBs, AOFBs and gliders has been deployed.
- Canadian MED online wave database: includes wave observations from 50 buoys since 1970, with 18559 days of measurement (statistics as at 24 August 2019)
- Surface wave measurement by drifting wave buoys (MR18-05C) [A20190702-001] ADS
- Waves in Ice (WII) drifting type wave buoy measurements in the Beaufort Sea for September and October 2016 [A20180306-001] ADS
- The 2010 Met Buoy Data, Klondike Study Area, NE Chukchi Sea dataset, data identifier: doi:10.5065/D60863BP: containing meteorological data collected for ConocoPhillips and Statoil on two buoys moored in the North-eastern Chukchi Sea during the open water season, July to September 2010. Met buoy data parameters include wind direction, wind speed, maximum wind speed, air temperature, relative humidity, air pressure, water temperature, significant wave and maximum wave heights, wave period, current speed and direction, and the voltage of the battery used to operate the meteorological instrumentation on the buoy. These data are part of a multi-year baseline environmental studies data set collected as part of the permitting process prior to drilling exploration wells. This dataset is part of the Pacific Marine Arctic Regional Synthesis (PacMARS) Project. This dataset contains the original CSV files, metadata files, and shapefiles in ZIP files.

- GTS wave data: there exist buoy data in GTS, e.g. WMO stations 48211, 48213, 48214 in the Beaufort–Chukchi seas
- SWIFT (Surface Wave Instrument Float with Track): including SWIFT data in MIZ program and in SeaState2015. The data can be downloaded in http://faculty.washington.edu/jmt3rd/SWIFTdata/ArcticOcean/.

3.3.1.4 Sea ice

Major parameters of sea ice relevant to Copernicus services are ice concentration, ice thickness, ice category, ice surface temperature (IST), ice drift and snow depth and temperature over sea ice.

- Ice drift derived from IABP ice drifting buoys (NRT and historical data, INS DAC)
- Unified Sea Ice Thickness Climate Data Record, 1947 onward, NSIDC (National Snow and Ice Data Centre): a concerted effort to collect as many observations as possible of Arctic and Antarctic sea ice draft, freeboard, and thickness and to format them consistently with clear documentation, allowing the scientific community to better utilize what is now a considerable body of observations. Parameters: Sea Ice thickness, Sea Ice Draft and Snow Temperature
- DMI ocean-ice database: at DMI, a database of in situ observations has been set up for use in the EUMETSAT OSISAF and other projects, as shown in Table 3.3.1. The data are available upon request. It should be noted that this database is partly overlapping with CMEMS INSTAC database. Parameters include IST, snow depth, radiation, ice thickness, surface temperature, sea ice drift and sea ice concentration etc.
- SEDNA (Sea ice Experiment - Dynamic Nature of the Arctic) dataset: can be obtained from NSIDC, observations on ice distribution, ice deformation, ice motion, sea ice dynamics, sea ice mass balance in 2007.
- Coordinated Eastern Arctic Experiment (CEAREX) Data: a multi-platform field program conducted in the Norwegian Seas and Greenland north to Svalbard from September 1988 through May 1989. Canada, Denmark, France, Norway and the United States participated in the experiment. Providing data on a full set of Meteo-ocean-ice-biogeochemical variables.
- Data from national Arctic expeditions: examples from Japan ADS include:
    - First Ice Camp of east of the Northern Northwind Ridge site at 50km northwest from Araon in by 2014 Araon Cruse. [A20141024-002]ADS
    - First Ice Camp of east of the Northern Northwind Ridge site at nearby Araon in by 2014 Araon Cruse. [A20141024-001] ADS
    - Second Ice Camp of West of the Northern Chukchi Plateau at near in Araon in by 2014 Araon Cruse. [A20141024-003] ADS
    - High resolution data set of sea ice production in the Arctic Ocean [A20141120-002] ADS
    - High resolution dataset of thin sea ice thickness in the Arctic Ocean [A20141120-001] ADS
    - Mooring data of sea ice and ocean off Barrow, Alaska during 2013-15 and 2014-15 [A20150917-001] ADS
- Data from research projects: there are a large number of datasets including sea ice parameters in the NSF ADC database, e.g., 104 IMB datasets are found.

Table 3.3.1 Database of Northern Hemisphere sea-ice measurements collected at DMI as part of an OSISAF project. Data are collected for internal use but will be available upon request. TA - atmospheric temperature, IST - ice surface temperature

	Type	Number	Period	Parameters	Data link
AWI	Multiple	147	2013 - present	TA, IST, Pressure, snow depth	http://data.meereisportal.de
MET-No	AWS on ship	1	2014 - 2015	Radiation	N/A
CRREL	Buoy (IMB)	108	2002 - present	TA, Pressure, snow depth, ice thickness, surface temperature	http://imb-crrel-dartmouth.org/live-data/
ECMWF	Multiple	2367	Up to present	TA, IST, Dew point temperature, air pressure, wind speed and direction	N/A
IABP	Buoy/drifting	2468	1991 - present	Sea ice drift	N/A
ITP	Buoy/drifting	110		Sea ice drift	N/A
IceBridge	Aircraft measurements	130	2009 - present	IST, Sea ice concentration, aircraft altitude, solar zenith angle	N/A

3.3.1.5 Biogeochemistry

Biogeochemical data (Dissolved oxygen, chl-a, nutrients, pCO2, PH, plankton etc.) can be obtained from following sources:

- International data centres: SeaDataNet, ICES, INSTAC, NODC (WOD and GLODAP)
- National data centres in Norway, Canada, French, UK, Poland, Finland, Sweden, Russia (restricted), China, Korean and Japan etc.
- BioArgo
- Research databases: e.g. NSF Arctic Data Centre, a variety of mooring monitoring programs

SeaDataNet and WOD

Since the sources of biogeochemical data in SeaDataNet and WOD 2018 are different, two datasets are highly complementary. In SeaDataNet, the amount of data in non-Nordic waters is negligibly small in comparison to that in the Nordic Seas: chl-a, nitrate and phosphate data in former only count for 1-3% of the latter (Tab. 3.3.2). In WOD2018, however, the ratio is 26-85%.

By assuming an overlapping rate of 0% or 100% between the two datasets, the number of profiles per annum is between 1223-1909, 1144-1731 and 1183-1736 for chl-a, nitrate and phosphate, respectively in the Nordic Seas; and 583-620, 198-213 and 144-153 for the non-Nordic Seas.

Table 3.3.2 The amount of biogeochemical data in SeaDataNet and WOD2018 during 2000-2017

	Chl-a		Nitrate		Phosphate	
	Non-Nordic Sea	Nordic Sea (60W-60E)	Non-Nordic Sea	Nordic Sea (60W-60E)	Non-Nordic Sea	Nordic Sea (60W-60E)
SeaDataNet	657	22020	271	20600	156	21310
WOD2018	10493	12334	3565	10552	2600	9929

Station distribution of nitrate observations in 2000-2017 in WOD 2018 and SeaDataNet is displayed in Figure 3.3.11

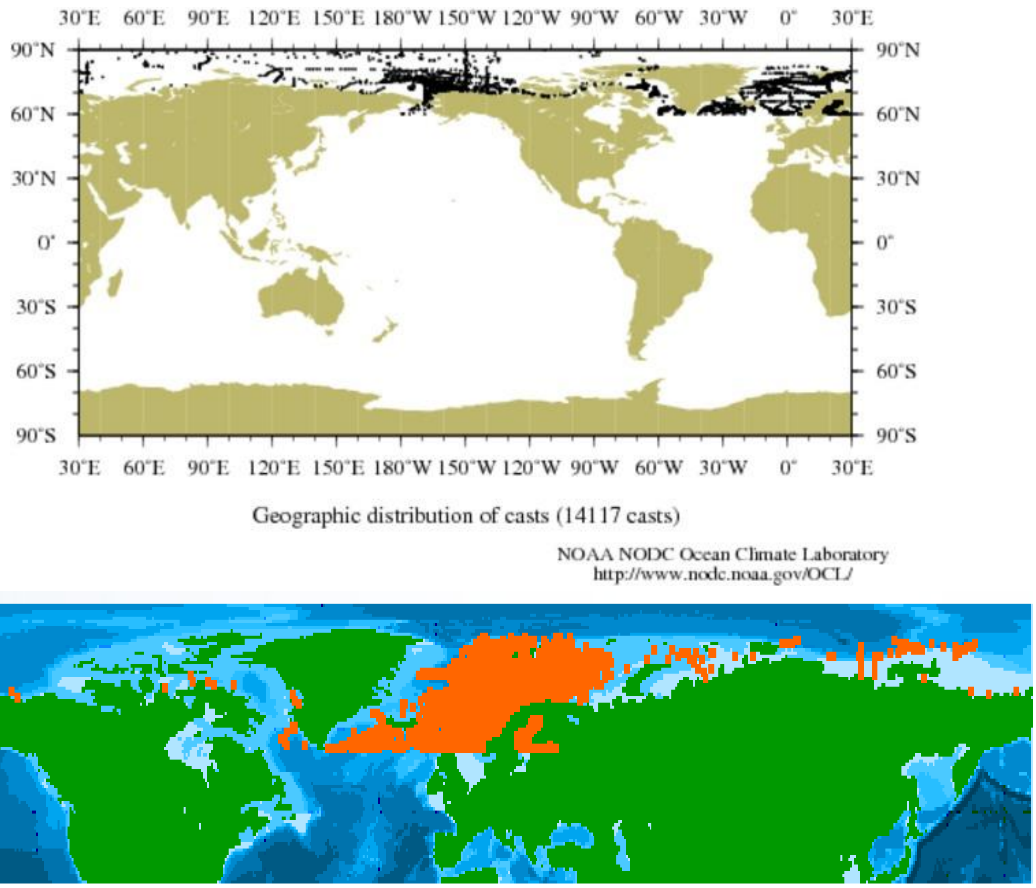

Figure 3.3.11 Station distribution of nitrate observations in 2000-2017 in WOD 2018 (upper panel) and SeaDataNet (lower panel).

ICES, GLODAPv2 and NMDC

It is expected that the ICES dataset has a big overlap with WOD and that the NMDC dataset has a big overlap with SeaDataNet. GLODAP (Global Ocean Data Analysis Project) is a uniformly calibrated open ocean data product on inorganic carbon and carbon-relevant variables. The station distribution of the three datasets is shown in Figure 3.3.12. It is obvious that GLODAP data is complementary to the other two datasets.

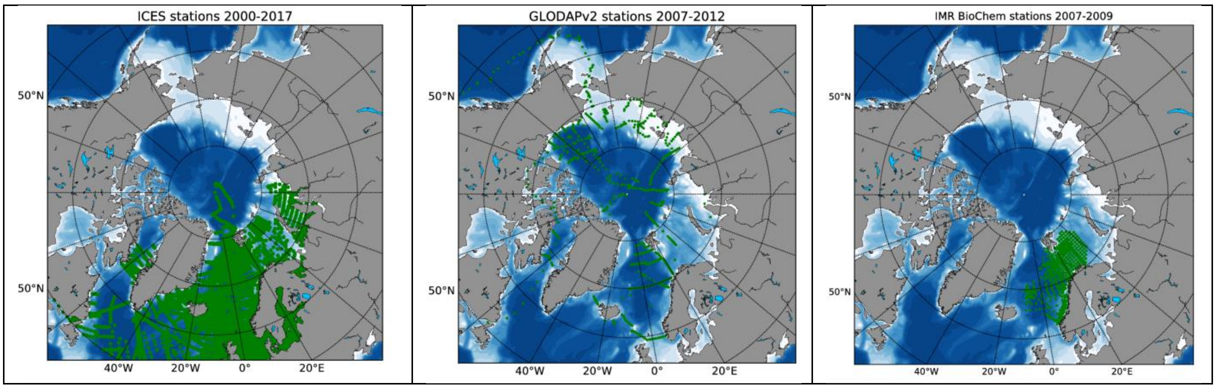

Figure 3.3.12 Nutrients profiles density from three databases: ICES, GLODAPv2 and NMDC (IMR) accessed in September 2017. Note the temporal coverage is unequal (source: Mercator Ocean et al. 2018)

BioArgo floats

Argo float maps and statistics can be found in http://www.jcommops.org/ftp/Argo/Maps/. The coverage of BioArgo in July 2019 is shown in Figure 3.3.13. It shows that there are about 13 BioArgo floats to the north of 60°N, which measure biogeochemical parameters in the Arctic. Most of them are found in the Nordic Seas. Such amount of BioArgo floats can generate about 468 profiles per annum.

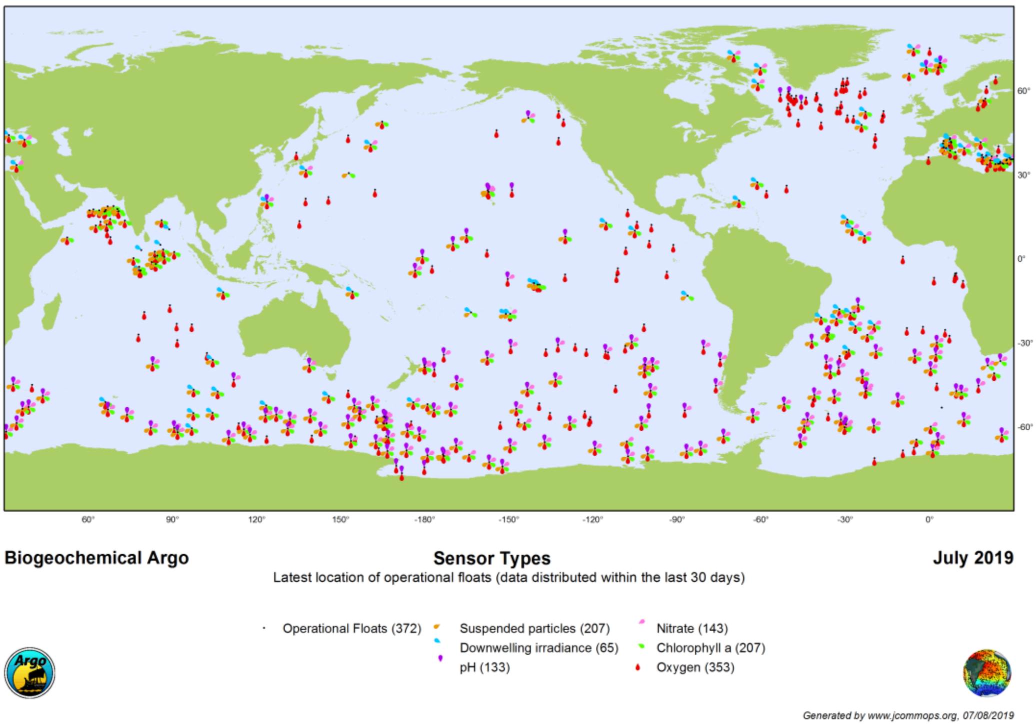

Figure 3.3.13 Argo float statistics: distribution of BioArgo floats in July 2019 (map source: JCOMM)

Other national data centres: here we will not go to an exhaustive assessment on all centres but only take Canada and Japan as examples.

***Canada**:* Canada MED section maintains a Biochemical database. The data can be accessed upon request at http://www.dfo-mpo.gc.ca/science/data-donnees/biochem/index-eng.html.

Japan: Arctic biogeochemical measurement can be found in Japan Arctic Data Service (ADS), example of data entries can be given here:

- Nutrients, carbon, chlorophyll a, $_{18}$O and CTD data in Bowdoin Fjord, Greenland in July 2016 [A20170420-002] ADS
- Nutrients, carbon, chlorophyll a, $_{18}$O and CTD data in Bowdoin Fjord, Greenland in July 2017 [A20190517-007] ADS
- Nutrients, trace metal, chlorophyll a, $_{18}$O, salinity organic matter and CTD data in Inglefield Fjord, Greenland in August 2018 [A20181001-002] ADS
- Optical data in the Bering and Chukchi Seas: T/S Oshoro-Maru cruise (C255) 2013 [A20131007-002] ADS
- Physical and chemical data during Umitaka-Maru cruises of the 58th Japanese Antarctic Research Expedition in January 2017[A20181220-001] ADS
- Rare earth elements, nutrient, oxygen isotope and strontium isotope samples in the Arctic Ocean of T/S Oshoro-Maru cruise (C255) 2013[A20131121-003] ADS
- Spatial distribution of dissolved organic matter in the Arctic Ocean and Bering Sea during T/S Oshoro-Maru cruise 2013 [A20130927-003]ADS
- Total inorganic carbon, total alkalinity and 18O data during the T/S Oshoro-maru cruise (cruise name: C255) in the Bering and Chukchi Seas in 2013. [A20130927-001] ADS

Research databases: there are a large number of datasets containing biogeochemical parameters such as:

- NSF Arctic Data Centre (many project datasets)
- Moorings and CTD in Coastal Greenland – data in http://data.g-e-m.dk
- Ecological monitoring in Young Sound – data in www.seanoe.org
- Moorings north of Svalbard in 2017 and 7 moorings in 2018 – French data in www.seanoe.org, polish data in IOPAN but they do not have a data management system (they consider PANGAEA to get a doi). During cruises other data are collected CTD, biogeochemical variables, turbulence
- Acidification mooring in the Fram Strait – data in Pangaea
- Carbonate chemistry in Kongsfjorden at Svalbard
- AREX
- ICOS high latitude monitoring
- ICOS high latitude monitoring (https://www.icos-cp.eu/)
- PAG (Pacific Arctic Group) and DBO (Distributed Biological Observatory): there will be 11 PAG and DBO research cruises in the Arctic region (details in Fig. 3.3.14), which are expected to monitoring biogeochemical observations.

Surface biogeochemical parameters are also measured by the Ferrybox, as described in 3.3.1.1.

2019 PAG and DBO Field Season (version 02_01_19): Sampling Contributors. Projects Key: AON=US Arctic Observing Network (National Science Foundation); ArCS=Arctic Challenge for Sustainability; ArcticEIS2=Arctic Ecosystem Integrated Survey, C30=Canada's Three Oceans; CHINARE=Chinese Arctic Research Expedition; DBO=Distributed Biological Observatory; EcoFOCI= JAMSTEC= Japan Agency for Marine-Earth Science and Technology; JOIS=KOPRI = Korea Polar Research Institute; MOSAiC= Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC); NIPR = National Institute of Polar Research; NOAA=National Oceanic and Atmospheric Administration; Office of Naval Research (ONR) Marginal Ice Zone (MIZ) project; PMEL=Pacific Marine Environmental Laboratory. **DBO Region Key**: DBO1=So. St. Lawrence Is., DBO2=Chirikov Basin, DBO3=So Chukchi Sea, DBO4=NE Chukchi Sea, DBO5=Barrow Canyon, DBO6=East Beaufort Sea, DBO7-Beaufort Sea Central, DBO8=Bathurst polynya region.

Dates 2019 (Port calls)	Ship	DBO Region	Projects	PAG contact	Chief Scientist
July 12-24 (Dutch Harbor-Utqiaġvik)	Sir Wilfrid Laurier	1,2,3,4,5	C30/DBO (AON)	Jackie Grebmeier jgrebmei@umces.edu	John Nelson John.Nelson@dfo-mpo.gc.ca
June-Sept (Shanghai-Shanghai)	Xuelong	-	CHINARE/MoSAIC	Jianfeng He hejianfeng@pric.org.cn	Jianfeng He hejianfeng@pric.org.cn
Aug 1-Oct 2 (Dutch Harbor-Nome-Nome-Nome-Dutch Harbor)	R/V Ocean Starr	2,3,4,5	Arctic IES (Integrated Ecosystem Survey)	Ed.Farley@noaa.gov	Geoff Lebon, Geoff.t.lebon@noaa.gov Ed Farley, ed.farley@noaa.gov Kris Cieciel, Kristin.cieciel@noaa.gov
Aug 1-25 draft (Dutch Harbor-Utqiaġvik)	Araon	1,2,3	K-AOOS (Korea-Arctic Ocean Observing System	Sung-Ho Kang shkang@kopri.re.kr	Eun Jin Yang ejyang@kopri.re.kr
Aug 27 – Sept 20 (Dutch Harbor-Nome-Nome-Dutch Harbor)	F/V Northwest Explorer	2	Northern Bering Sea Assessment	Ed.Farley@noaa.gov	Jim Murphy, jim.murphy@noaa.gov
Aug 2-23 (Utqiaġvik - Utqiaġvik)	Healy	1,2,3,4,5	DBO/NCIS=Northern Chukchi Integrated System	Jackie Grebmeier jgrebmei@umces.edu	Robert Pickart rpickart@whoi.edu and Jackie Grebmeier jgrebmei@umces.edu
Sept (Nome-Nome)?	Norseman II	3	Bering Strait Mooring Project/AON	Rebecca Woodgate woodgate@apl.washington.edu	Rebecca Woodgate woodgate@apl.washington.edu
Sept 18-Oct 6 (Dutch Harbor-Kodiak)	Dyson	1 and M8	EcoFOCI	Phyllis Stabeno, Phyllis.stabeno@noaa.gov	Geoff Lebon geoffrey.t.lebon@noaa.gov
Sept -Oct?	Louis S. St-Laurent	-	JOIS/AON-BGOS	Bill.Williams@dfo-mpo.gc.ca	Bill.Williams@dfo-mpo.gc.ca
Oct?	Sir Wilfrid Laurier	4,8	C30	Bill.Williams@dfo-mpo.gc.ca	Humfrey.Melling@dfo-mpo.gc.ca
Sep 27- 10 Nov 2019 (Sekinehama, Japan, return Sekinehama, Japan)	Mirai	1,2,3	Japanese Atmospheric cruise; National Institute of Polar Research (NIPR)	Takashi Kikuchi takashik@jamstec.go.jp	Dr. Kazutoshi Sato satokazu@mail.kitami-it.ac.jp

Figure 3.3.14 2019 PAG and DBO Cruise Plan Table (03-14-19)

3.3.1.6 Data identified by INTAROS

In the H2020 project INTAROS, a deliverable was made on "Report on present observing capacities and gaps: ocean and sea ice observing system" (INTAROS,2018). An Arctic data survey was carried out in 2017, mainly based on current capacities in INTAROS partners. The survey was made for a much wider purpose than Copernicus Service.

Although only monitoring activities with the consortium partners were evaluated, significant amounts of physical and biogeochemical observations have been identified, which potentially can be used for CMEMS applications. It should be noted that these extra amounts of data covering mainly Nordic Seas and entire Barents Sea. Areas with less sampling are the Greenlandic and Icelandic offshore and open waters. INTAROS has got funding to perform some in situ observations – some systems were deployed in 2017 but most during summer 2018:

- Moorings in Young Sound data in http://data.g-e-m.dk
- Moorings and CTD in Coastal Greenland – data in http://data.g-e-m.dk
- Ecological monitoring in Young Sound – data in www.seanoe.org
- Moorings north of Svalbard in 2017 and 7 moorings in 2018 – French data in www.seanoe.org, polish data in IOPAN but they do not have a data management system (they consider PANGAEA to get a doi). During cruises other data are collected CTD, biogeochemical variables, turbulence
- Acidification mooring in the Fram Strait – data in PANGAEA
- Carbonate chemistry in Kongsfjorden at Svalbard
- Underwater acoustics in Kongsfjorden
- Ice-tethered platforms in Central Arctic – unfortunately only worked for one month.

- Snow and Ice mass balance Arrays (SIMBA) central Arctic
- Glider lines along Atlantic Water Pathway in Fram Strait – data in www.seanoe.org
- Ferrybox lines in Barents Sea – available on NIVA ftp server
- BioArgo floats in Baffin Bay – data in Coriolis

More details are given from the INTAROS report D2.1 (INTAROS, 2018), summarized as below:

NERSC: acoustic measurements are collected in several projects during 2008 – 2016 from Fram Strait Multipurpose Acoustic system (2008 -2014) and Canada Basin Acoustic Propagation Experiment (CANAPE 2016). Parameters observed are acoustic travel time and ambient noise. Both kind of data are included in the INTAROS data catalogue. Derived parameters are averaged sound speed through inversion of acoustic travel times and depth-range temperature along acoustic thermometry sections.

UIB (University of Bergen): At station Mike and the area north of Svalbard, ocean buoys and moorings instruments have been providing observed parameter(s): Surface air temperature, salinity, temperature, salinity, pressure, currents, fCO2, NO3 at selected depths on the mooring. Located in Norwegian Sea the moorings have both real time and delayed mode capabilities. The site represents the longest existing homogeneous time series from the Deep Ocean (1982-2018). UIB have been the major supplier of carbon system data since 1987 and can today provide all carbon system variables including stable isotopes. The group has had several projects in the Greenland Sea, Storfjorden, Irminger Sea, Iceland Sea, Barents Sea and in the Nordic Seas and North Atlantic in general. ICOS Carbon Portal (CP): https://www.icos-cp.eu/).

IMR (Institute of Marine Research, Norway):

- An A-TWAIN mooring array has been operated since 2012 in north of Svalbard (81-82° N, 31° E), with 2-3 ocean moorings and turnover every 1-2 years and observed parameter(s) on water velocity and T/S profiles. Data are available through the Norwegian Polar Data Centre at the Norwegian Polar Institute (https://data.npolar.no/home/).
- A Barents Sea Opening Mooring Array has been operated since 1997 with 5 fixed moorings along a section across the western Barents Sea, with observed parameters of surface and subsurface ocean temperature and current velocity. Data is recorded every 20 minutes. The data are handled by the Norwegian Marine Data Centre, Bergen, Norway. Available on supervised request through originator. Most data are also shared with and stored at the ICES.
- An IMR-PINRO Ecosystem Survey has been operated since 2004, running annually in August-September in the Barents Sea (Norwegian, Russian and international sectors, 68-82°N, 5-60°E). The survey is based upon in situ measurements from scientific vessels (normally three). The following measurements are routinely made (additional measurements may be done sporadically): T/S, nutrients, primary production, secondary production, abundance, size, age, distribution of many fish stocks, etc. The data are handled by the Norwegian Marine Data Centre, Bergen, Norway. Most data are also shared with and stored at the International Council for the Exploration of the Seas (ICES) in Copenhagen, Denmark. Records are updated with new observations irregularly, but roughly annually. Data is available on supervised request through originator.
- A Barents Sea Winter Survey has been operated since 1876, running annually in January-February in the Barents Sea (Norwegian, Russian and international sectors, 68-80° N, 7-56

°E). Measured parameters and data management are similar to the IMR-PINRO Ecosystem Survey in summer.

- A Fixed hydrographic (near coastal) station network has been operated since 1936, at 3 stations in northern Norwegian waters: Ingøy (71.13 °N, 24.016°E), Eggum (68.367° N 13.633°E) and Skrova (68.116 °N, 14.533° E) with T/S monitoring 3-4 times per month at Skrova and 1-2 times per month at Ingøy and Eggum.
- 6 fixed hydrographic sections have been operated in Norwegian Seas with measurements on T/S, nutrients, primary production and secondary primary production. Hydrographic sections are a compromise between observing at a few positions with high frequency, and undertaking a spatially broader coverage with many measurement points with low frequency.
- A vessel mounted ADCP system was operated annually in an area (77.5N 3E), (77.5N 25E), (82.5N 3E), (82.5N 25E) during August 2014 to 2018 with possible continuation. Surface and subsurface current velocities were measured using a research vessel mounted ADCP system Instruments.

AWI (Alfred Wegener Institute): AWI contributes from the FRAM observing system, oceanographic and biogeochemical datasets from an area within the Fram Strait between 78- and 80.1°N and 5.5°W and 11.1°E:

- *HAUSGARTEN observatory* displays 21 permanent stations covering a water depth range of 250 to 5500 m water depth. Repeated sampling and the deployment of moorings and different free-falling systems, which act as observation platforms has taken place since the beginning of the station in summer 1999;
- Data from the AWI 79°N *Fram Strait Mooring Array* contributes with temperature, salinity, current velocity, oxygen;
- *CTD Measurements by RV Polarstern in Fram Strait contributes* 45 cruise legs with more than 2500 CTD profiles starting in 1987 and are available from the PANGAEA repository;
- *Vessel-Mounted ADCP on RV Polarstern* measures a profile of the ocean current speed in the upper 320 m of the water column. The data flow, quality checking and archiving is not very mature. For the FRAM domain data from only 10 cruises (between 1993 and 2017) are available from the PANGAEA data repository.

IOPAN (The Institute Of Oceanology Of The Polish Academy Of Sciences) : has a Long-term large-scale multidisciplinary Arctic monitoring program /AREX) and Long-term Monitoring in Svalbard Fjords.

- AREX has operated annually repeated summer survey of 2-month duration in the eastern Nordic Seas and Fram Strait (70-81°N; 0-22°E) since 1988. T/S and current profiles are measured. Raw and processed measurements are stored in the IOPAN database (under development) and provided on request. Data are publicly available with the protection period of 2 years. It is also planned to submit the main CTD dataset to the PANGAEA (https://www.pangaea.de/), a world data centre and a member of the International Council for Science (ICSU) World Data System.
- The Long-term Monitoring in Svalbard Fjords covers Hornsund (76°50’ - 77°10’N, 15-17°E), Kongsfjorden (78°50’-79°05’N, 11- 13°E) and shelf region next to the fjords’ outlets. The monitoring started from 1995 and still ongoing with annually repeated summer surveys and

spring and autumn measurements in selected years. CTD profiles are measured. Data are publicly available with the protection period of 2 years.

DTU (Technical University Of Denmark):

- *IOC Tide Gauge Network for Greenland:* DTU provide daily data and high-frequency data (5 minutes) at 4 tide gauge stations in Greenland. Sensors/instruments include tide gauges (relative sea level), thermometer, barometer and GPS receiver. Data available for download: http://www.ioc-sealevelmonitoring.org

AU (Aarhus University):

- *Greenland Ecosystem Monitoring Programme:* The system for physical ocean data consists of a fixed mooring combined with a repeated CTD transect every year in August since 2003 (74 N 21 W) - (64°N; 51°W). 2003- ongoing. Data are quality checked and available at the Greenland Ecosystem database (GEM database), see http://data.g-e-m.dk/

NIVA (Norwegian Institute For Water Research)

- *Barents Sea Ferrybox system: along a ferry section* between Tromsø, Norway and Longyearbyen, Svalbard, the system measures T, S, chl-a fluoresces, turbidity, oxygen, pCO_2 and pH at 5-meter depth. The data is already available in CMEMS but hasn't been used for modelling applications.

UNIS (The University Centre in Svalbard)

- *UNIS ocean observing system around Svalbard:* The UNIS ocean observing system comprises fixed moorings carrying oceanographic instruments e.g. current meters, CTD, pressure sensors, covering waters surrounding Svalbard. It has been operated since 2003.

3.3.1.7 Data usage in ARC MFC and satellite TACs

Arctic in situ data are collected by CMEMS INSTAC. Part of the collected data have been used in both TACs and MFCs. Based on the survey and CMEMS Product User Mannuals and QUID reports, a summary of the usage of in situ data is given in Table 3.3.3a for TACs and Table 3.3.3b for Arctic MFC, corresponding to each product. The CMEMS product list covers a wide range of parameters: sea level, water temperature, salinity, currents, sea ice variables, wave variables, dissolved oxygen, nutrients, pH, carbon related variables, chl-a, ammonium, primary production, phytoplankton and zooplankton etc.

For TACs, in situ observations are used mainly for calibration/validation (cal/val) of Level 3 (L3) and Level 4 (L4) products. For some secondary products, e.g. ice drift, in situ observations can also be used for developing retrieving algorithms. Currently only Sea Level (SL) -, Sea Surface Temperature (SST)- and Ocean Sea Ice (OSI) TACs use Arctic in situ observations. No Arctic in situ data are used in OC/Wind/Wave TACs.

Temperature, salinity and biogeochemical profiles:

For the profile data, the amount of data assimilated for reanalysis is about 4000-5000 T/S profiles per year, mainly from Argo, ITP and R/Vs during the period 1993-2016. For validation purpose, 300 chl-a profiles from BioArgo floats, which are special Argo floats mounted with biochemical measurement devices, are used for model validation.

SST from surface drifters: they provide near real time (NRT), high quality SST measurements which are used in CMEMS for validating NRT forecast.

Surface currents: they are mainly calculated from drifting buoys (with drogue at 15m). There are HF radars in Northern Norway but these are better assimilated by downstream services. Their coverage is not worth the effort for a large area like the Arctic. The drifter-derived surface currents are now used in the validation of NRT forecasts.

Current profiles: they are taken from NABOS (Nansen and Amundsen Basins Observational System) moorings from University of Alaska Fairbanks (Microwave Temperature Profilers - MMTP), IMR moorings in the Barents Sea Opening and Bering Strait moorings from Applied Physics Lab. in University of Washington (APL-UW). The data are used for validating physical reanalysis.

Ice drift: they are taken from the IABP (International Arctic Buoy Program) drifting buoys and used for validating the ice drift acceleration and its seasonal cycle in the physical reanalysis

Waves (Significant wave heights, peak period, mean direction): they are taken from moored wave buoys (Only from Norwegian oil platforms and on Iceland) and used in wave model calibration and validation, only a short period is covered.

Table 3.3.3a Current use of Arctic in situ observations in CMEMS Thematic Assembly Centres (TACS)

CMEMS users	**CMEMS Product name**	**Current usage of Arctic in situ observations**										
		Purpose of use			**Data characteristics**							
		AD	**CV**	**DA**	**Dataset name**	**Provider**	**Para-meter**	**No. of stations**	**Cover-age**	**update freq.**	**Time-liness**	**Period**
SL TAC	SEALEVEL_ARC_PHY_L3_NRT_OBSERVATIONS_008_038		√		GLOSS/CLIVAR & PSMSL	UHSLC	sea level	~10	Arctic	monthly	N/A	N/A
	SEALEVEL_GLO_PHY_L4_NRT_OBSERVATIONS_008_046		√			UHSLC	sea level	~100	Global	monthly	N/A	N/A
	SEALEVEL_GLO_PHY_L4_REP_OBSERVATIONS_008_047		√		GLOSS/CLIVAR & PSMSL	UHSLC	sea level	~100	Global	monthly	N/A	N/A
	SEALEVEL_GLO_PHY_L3_REP_OBSERVATIONS_008_045		√			UHSLC	sea level	~100	Global	monthly	N/A	N/A
OC TAC	OCEANCOLOUR_ARC_CHL_L4_NRT_OBSERVATIONS_009_087		√		No data used	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	OCEANCOLOUR_ARC_CHL_L3_NRT_OBSERVATIONS_009_047, L3/L4 Reprocessed products		√			N/A	N/A	N/A	N/A	N/A	N/A	N/A
SST TAC	SEAICE_ARC_SEAICE_L4_NRT_OBSERVATIONS_011_008; L4_NRT_OBSERVATIONS_011_003; L3_REP_OBSERVATIONS_011_010		√		Drifters	ICOADS	SST	~132	Arctic Ocean	hourly	NRT	Hourly
			√		IMB buoy	DMI	IST	~8		hourly	offline	2012
			√		Argo T/S	INSTAC	SST	~80		10 days	NRT	10 days
SIC TAC			√		Drifter data	ICOADS	SID	~132	Arctic Ocean	hourly	NRT	
Wind TAC	WIND_GLO_WIND_L3_NRT_OBSERVATIONS_012_002		√		No data used	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Wave TAC	WAVE_GLO_WAV_L3_SWH_NRT_OBSERVATIONS_014_001		√		No data used	N/A	N/A	N/A	N/A	N/A	N/A	N/A

ARC: Arctic: SL: Seal level; OC-Ocean color; REP: Reprocessed; BGC: Biogeochemical; PHYS: Physics; INS: In situ
AD: Algorithm development; CV: Calibration/Validation; DA: Data assimilation; NRT: Near real time; RAN: reanalysis;

Table 3.3.3b Current use of in situ observations in CMEMS Marine Monitoring and Forecasting Centres (MFCs)

CMEMS users	**CMEMS Product name**	**Current usage of Arctic in situ observations**										
		Purpose of use			**Data characteristics**							
		AD	**CV**	**DA**	**Dataset name**	**Provider**	**Para-meter**	**No. of stations**	**Cover-age**	**update_ freq.**	**Time-liness**	**Period**
ARC MFC	ARCTIC_ANALYSIS_FORECAST_PHYS_002_001_A	√	√	√	Argo & R/Vs TS	INSTAC	T/S	~2000/y	Nordic Seas	daily	NRT & offline	2010-
			√		Drifter buoy	INSTAC	SST	~132	Arctic Ocean	hourly	NRT	2012-
			√		Drifter buoy	INSTAC	U,V	~132		hourly	NRT	2012-
		√			Sea ice drift buoy	IABP-SIDFEx	UICE,VICE	~10 at any time		3hourly	NRT	1980-
			√	√	ITP-TS	INSTAC	T/S	~400/y		daily	NRT	2006-
	ARCTIC_ANALYSIS_FORECAST_BIO_002_004	√			R/V BGC	INSTAC	chl-a, N	~30	Nordic Seas	daily	offline	2007-2008
			√		Bio-Argo	INSTAC	chl-a, O2	~300	Irminger Sea	daily	offline	2015-
	ARCTIC_REANALYSIS_PHYS_002_003		√	√	Argo, ITP & R/Vs TS	INSTAC	T/S	~3000/y	Arctic Ocean	daily	offline	2003-2017
			√	√	ICES Ocean. Data	ICES	T,S	~1000/y	Nordic Sea	daily	offline	1991-2009
			√	√	R/Vs	AARI, IOPAS, MMBI	T/S	~1000/y	Nordic Sea	daily	offline	1991-2009
			√		IMB buoys	CRREL Portal[1]	SIT	3	Beaufort Sea	daily	offline	2013-2016
			√		BGEP moorings	WHOI	SIT	4	Beaufort Sea	daily	offline	2013-2016
			√		Drifter buoy	INSTAC	SST	~132	Arctic Ocean	hourly	NRT	1991-
			√		Drifter buoy	INSTAC	U,V	~132		hourly	NRT	1991-
			√		IABP buoys (sea ice)	NSIDC	UICE,VICE	~10 at any time		3hourly	NRT	1980-
			√		Tide Gauges	PSMSL	sea level	42	Arctic Ocean	monthly	offline	1991-2009
	ARCTIC_REANALYSIS_BIO_002_005		√		R/V data	INSTAC	chl-a, N	~30	Nordic Seas	daily	offline	2007-2008
	ARCTIC_ANALYSIS_FORECAST_WAV_002_010		√		Platform-Buoy Wave	Met.no	Waves	~26	NOR-offshore	hourly	offline	3-4 months

[1] http://imb-crrel-dartmouth.org/imb.crrel/buoysum.htm

Sea level: they are measured with tide gauges. Sea Level TAC and ARC MFC use high frequency sea level data only from Norwegian tide gauges. Monthly historical data from 10 stations in GLOSS database and 42 stations in PSMSL database have been used for validating the physical reanalysis products and reprocessed satellite products.

Ice thickness: measured from airborne campaigns (Operation Ice Bridge), Drifting Ice Mass Balance Buoys (from CRREL, Dartmouth, USA) and moorings in the Beaufort Sea (Beaufort Gyre Experiment Programme, BGEP from WHOI, USA), they are used for the **v**alidation of the physical reanalysis products. The spatial coverage is very much limited.

Snow depth on ice: data are obtained from research campaigns and Integrated Arctic Ocean Observing System (iAOOS) buoys from LOCEAN. They are used for the validation of the physical reanalysis products. However, the data are too few.

By comparing with the available data listed in 3.3.1.1 – 3.3.1.6, it can be concluded that there is still a significant amount of Arctic in situ data that are accessible but not used in CMEMS MFCs and satellite TACs. For example, the number of T/S profiles from several big data centres SeaDataNet, ICES, WOD and EMODnet has reached about 20000 per year but the used number of profiles for ARC MFC reanalysis is only 4000-5000 T/S profiles per year. This is partly due to the fact that not all observations have been collected by INSTAC, and partly due to the quality and impacts of a certain data type e.g. APB data, has not been quantified for the reanalysis.

A summary of part of the existing observations is given in Table 3.3.4. The table also identifies their potential use in correspondent CMEMS TACs and ARC MFC. It should be noted that the table is not exhaustive; many datasets described in s3.3.1.1-3.3.1.6 have not been included.

Table 3.3.4 Observations exist but not used in CMEMS: DFO – Fisheries and Oceans Canada, DTU – Technical University of Denmark, iAOS – Integrated Arctic Observing System, ICES – International Centre for Exploring the Seas, IMR – Institute of Marine Research, INTAROS – Integrated Arctic Observing System, NIVA – Norwegian Institute for Water Research, NDBC – National Data Buoy Centre, NMA – Norwegian Maritime Authority, NODC – National Ocean Data Centre, REP – Reprocessed, TG – Tide Gauge, WOD – World Ocean Database,

CMEMS users	**Existing but not used data**									
	Data characteristics								**Purpose of use**	**fit4purpose**
	Data source	**Data provider**	**Parameter**	**No. of stations**	**area_ covered**	**update_ freq.**	**Delivery**	**Period**	**accessibility**		
SL TAC, ARC MFC	Norwegian TG data	NMA	sea level	~17	Norwegian coast	5-60 min.	real time	On going	open, free	NRT/REP/RAN product cal/val	NRT/REP/RAN
	DTU TG data	DTU		4	Greenland coast	.			open, free		NRT/REP/RAN
	Canadian TG data	DFO		3~10	Canadian coast				open, free		NRT/REP/RAN
	US TG data	NOAA		~10	Alaska coast				open, free		NRT/REP/RAN
	Russian TG data	Unknown		13	Russian coast				Restricted		NRT/REP/RAN
OC TAC, ARC MFC	ICES Oceano. Database	ICES	chl-a, profile	20-30 yr^{-1}	Nordic Seas	irregular	>1year	On going	open, free	NRT/REP or RAN product cal/val, assimilation	REP/RAN only
	iAOS/IMR cruises	INTAROS/IMR		>1000 yr^{-1}		irregular	Offline		Restricted		
	WOD2018	NODC		>1500 yr^{-1}		irregular	Offline		Open, free		
	Arctic BioArgo	Ifremer		Unknown		10day	Real time		open, free		NRT/REP/RAN
	iAOS Ferry	NIVA	Chl-a, 5m	1	Barents Sea	days	Real time		open, free		NRT/REP/RAN
SST TAC, ARC MFC	iAOS moorings	INTAROS	SST, T/S	>15	Nordic Seas	hourly	Real time	On-going	open, free	NRT/REP or RAN product cal/val, assimilation	NRT/REP/RAN
	iAOS ferrybox	NIVA	SST, SSS	1	Barents Sea	days	Real time		open, free		NRT/RAN

	ITP	WHOI	T/S,	>3	Arctic	hourly	Real time		open, free		NRT/REP/RAN
	iAOS/IMR cruises	INTAROS IMR	SST, T/S	>1000 yr^{-1}	Nordic Seas	Irregular	Offline		Partly restricted		REP/RAN only
	WOD2018	NODC	SST, T/S	>15000yr^{-1}	Arctic, not Russian side	Irregular	Offline		open, free		REP/RAN only
	UDASH	AWI	SST, T/S	8000 yr^{-1}	Arctic	Irregular	Offline	1980-2015	open, free		REP/RAN only
	ICES Oceano. Data	ICES	SST, T/S	>400 yr^{-1}	Icelandic/SE Greenlandic	Irregular	Offline	On-going	open, free		REP/RAN only
	Russian Oceano. Data	Russian	SST, T/S	Unknown	Arctic ice-free water on Russian side	Irregular	Offline		Restricted		REP/RAN only
OSI TAC, ARC MFC	IMB buoys	SIMBA	SIT, IST, snow depth,	Unknown	Central Arctic	Hourly	Real time	unknown	Open, free	NRT/REP or RAN product cal/val, assimilation	NRT/REP/RAN
		PMEL/CRREL		3			offline	2001-20013			
	Nord Pole Drifting data	Russian	Ice drift	Unknown	Central Arctic	Hourly	Real time	On going	Restricted		NRT/REP/RAN
	Russian Ice coverage	Russian	Ice coverage	Unknown	Arctic on Russian side	Unknown	Unknown				NRT/REP/RAN
	Russian coastal ice	Russian	Coastal ice	75	Russian coast	Unknown	Real time				NRT/REP/RAN
Wind TAC	Drifter buoy data	INSTAC	SSW	uncertain	Arctic Ocean	hourly	NRT	On-going	open, free	NRT/REP product cal/val	TBC
	SHIP data_US	NDBC	SSW	uncertain	Alaska waters	hourly	NRT		open, free		TBC
	Mooring buoy data_NOR	Met.no	SSW	26	Norwegian offshore	hourly	NRT		Restricted		TBC
	Mooring buoy data_US	NDBC	SSW	4	Alaska offshore	hourly	NRT		open, free		TBC
	Mooring buoy data_CAN	DFO	SSW	2	Canadian offshore	hourly	Offline	-2018	open, free		REP only, TBC

Wave TAC ARC MFC	SHIP data_US	NDBC	Hs, Dp	uncertain	Alaska waters	hourly	NRT	On-going	open, free	NRT/REP or RAN product cal/val,	TBC
	Mooring buoy data_NOR	Met.no	Hs, Tp etc	26	Norwegian offshore	hourly	NRT		Restricted		NRT/REP/RAN
	Mooring buoy data_US	NDBC	Hs, Tp etc	4	Alaska offshore	hourly	NRT		open, free		NRT/REP/RAN
	Mooring buoy data_CAN	DFO	Hs, Tp	2	Canadian offshore	hourly	Offline	-2018	open, free		REP/RAN
ARC MFC	ICES Oceano.	ICES	O2,nutrient	100 yr^{-1}	Nordic Seas	irregular	>1year	On going	open, free	NRT/REP or RAN product cal/val, assimilation	REP/RAN only
	iAOS/IMR cruises	INTAROS/IMR	O2, pCO2, pH, N, P, Si, PP	>1000 profiles yr^{-1}		irregular	Offline		Restricted		
	WOD2018	NODC	O2	3300 yr^{-1}	Arctic	irregular			open, free		
	WOD2018	NODC	N, P, Si	400-700 yr^{-1}	Nordic Seas	irregular	Offline		Open, free		
	Actic BioArgo	Ifremer	O2, turbidy	Unknown	Nordic Seas	10day	Real time		open, free		NRT/REP/RAN
	iAOS Ferry	NIVA	O2, pCO2, turbidity	1	Barents Sea	days	Real time		open, free		NRT/REP/RAN

3.3.15 Data with restricted availability

There are also some Arctic marine data with restricted availability, e.g. most of the recent years' data from Russia, cruise observations in Greenland and Iceland etc. In the INTAROS project, Russian partners (RIHMI-WDC: Russian Research institute of Hydrometeorological Information – World Data Centre, AARI: Arctic and Antarctic Research Institute and NIERSC St. Petersburg) have worked together to provide a survey of Russian metadata.

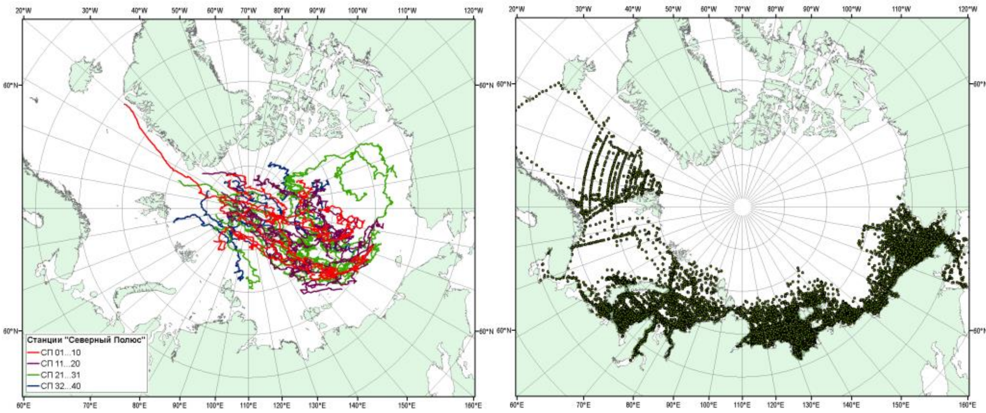

Figure 3.3.15 Left: "Nord Pole" drifting expedition stations (1937-2013). Right: oceanographic sampling stations from Russia

In January 2018, the Russian monitoring activities included 22 hydro coastal observation stations/posts, 13 sea level stations, research vessels expeditions with 4 ships and "North Pole" drifting expeditions (40). In addition, there are 75 coastal ice observation stations, river-sea ice monitoring and sea ice coverage monitoring activities. The station distribution is shown in Figure 3.3.15 and Figure 3.3.16.

The figures show that Russia has established a comprehensive and well covered monitoring network on physical parameters (sea level, hydrography and sea ice). It has not been possible in this study to document the extent of Russian monitoring programs of waves and biogeochemical parameters

The oceanographic data largely fill the gaps in ice-free waters on the Russian side. The "Nord Pole" drifting expedition data fill the ice-related measurements in Arctic ice-covered areas. However, these data are currently not available to CMEMS.

Monitoring cruises take place in Greenlandic waters and Icelandic waters several times a year by Greenland Institute for Natural Resources and Icelandic Marine and Fishery Research Institute. These data currently have restricted access. Icelandic sea level data are also restriced for access.

Finally it is believed that there exists an unquantifiable amount of ocean observations collected via research projects that are not visible in any data reposatories but are hidden on participating scientists' computers. These data are difficult to discover but could be very valuable to the community.

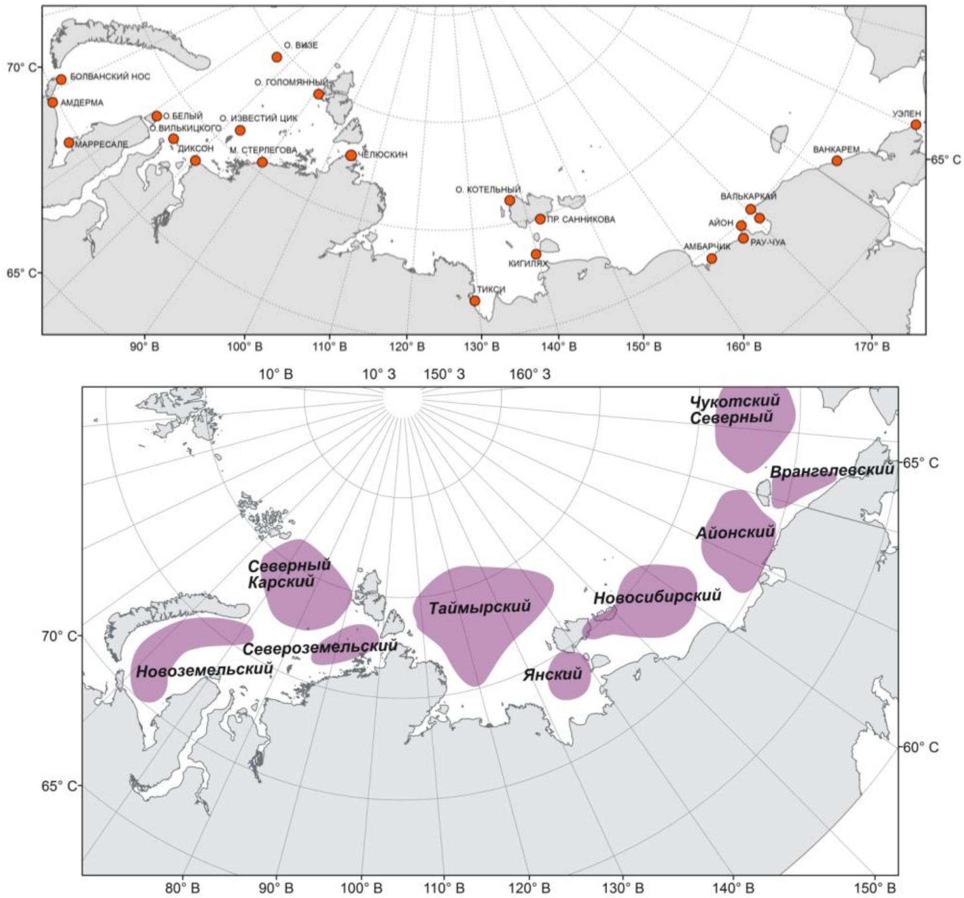

Figure 3.3.16 Upper panel: Russian coastal station network; Lower panel: Russian ice coverage monitoring.

3.4. Cryosphere

3.4.1. Freely available data

The Global Cryosphere Watch (GCW) is the WMO programme for monitoring the cryosphere. It is establishing a core network of surface observation sites, the CryoNet, which covers all parts of the cryosphere (ice sheets, ice shelves, glaciers, permafrost, snow, sea ice, river/lake ice) and also contributes to the standardization of best practices and easy access to data.

The CryoNet is built from existing stations. A station has to fulfil a set of requirements to be accepted as a CryoNet station:

- cryosphere measurements should be representative of the surrounding region,
- the responsible agency must be committed to sustain long-term operations,
- metadata should be kept up-to-date and made available for the GCW Portal,
- instruments, methods and QC should follow GCW regulations and best practices,
- data must be made freely available, in near real-time when possible.

Figure 3.4.1 Location of Arctic CryoNet stations

The GCW network also includes a number of contributing stations that do not fulfil all the CryoNet requirements. These stations may, however, provide valuable data and help close the gap related to the sparse coverage of data in the Arctic.

The GCW Data Portal is so far a pre-operational service. It consists of metadata from a number of data centres, but the actual data are still provided through the interfaces supported by the originating data centre.

3.4.1.1 Ice sheets (Greenland)

The Danish Meteorological Institute (DMI) has a network of meteorological stations around the margins of Greenland and the Summit Station at the top of Greenland ice sheet. These are available via GTS and may be downloaded at www.polarportal.dk

The Greenland Ice Sheet Monitoring Network (GLISN; www.glisn.info) consists of a network of 33 seismometers that can be used to detect glacial earthquakes and calving events, and seiche events resulting from iceberg calving.

3.4.1.2 Glaciers

Global Terrestrial Network for Glaciers (GTN-G) (https://www.gtn-g.ch) provides a browser for available glacier data. It is run by the World Glacier Monitoring Service (WGMS), the National Snow and Ice Data Centre (NSIDC) and the Global Land Ice Measurements from Space (GLIMS) and authorized under the Global Climate/Terrestrial Observing systems. All data submitted to GTN-G are considered public domain for non-commercial use and made digitally available through operational services at no cost. GTN-G data are open access under the requirement of correct citation (CC BY 4.0 creative commons license).

The World Glacier Monitoring Service (https://wgms.ch) provides standardized observations on changes in glacier mass, volume, area and length. All data and information is freely available for scientific and educational purposes but the user should acknowledgement WGMS and/or original investigators and sponsoring agencies according to the available meta-information.

GlacioBasis (A. P. Olsen and Zackenberg) contribute to the Global Terrestrial Network for Glaciers (GTN-G) through WGMS (http://g-e-m.dk/gem-science-programme/glaciobasis-programme/)

3.4.1.3 Permafrost

The Global Terrestrial Network for Permafrost (GTN-P, www.gtnp.org) is the main international programme concerned with monitoring permafrost parameters. GTN-P is developed under the Global Climate Observing System (GCOS) and the Global Terrestrial Observing Network (GTOS) to obtain a comprehensive view of changes in the active layer thickness and permafrost temperature. The map shows active boreholes.

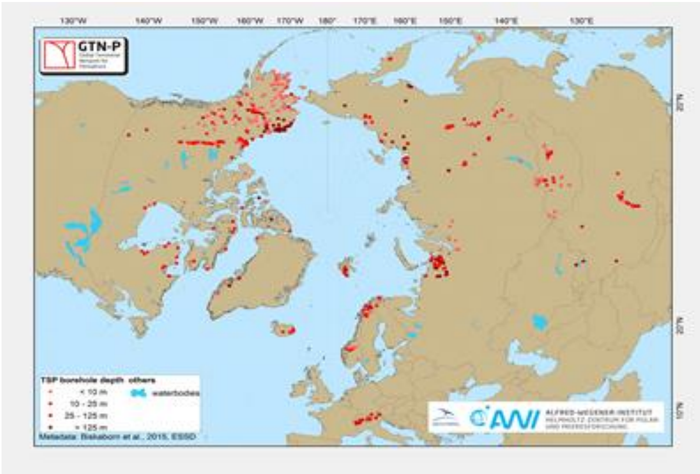

Figure 3.4.2 Borehole sites contained in the GTN-P Database

3.4.1.4 Snow

The WMO GTS system delivers ground-based snow depth measurements from SYNOP stations (e.g. used by the ECMWF data assimilation system). The number of observations pr. grid square is shown below. The plot illustrates the very sparse coverage of snow measurements from Arctic SYNOP stations.

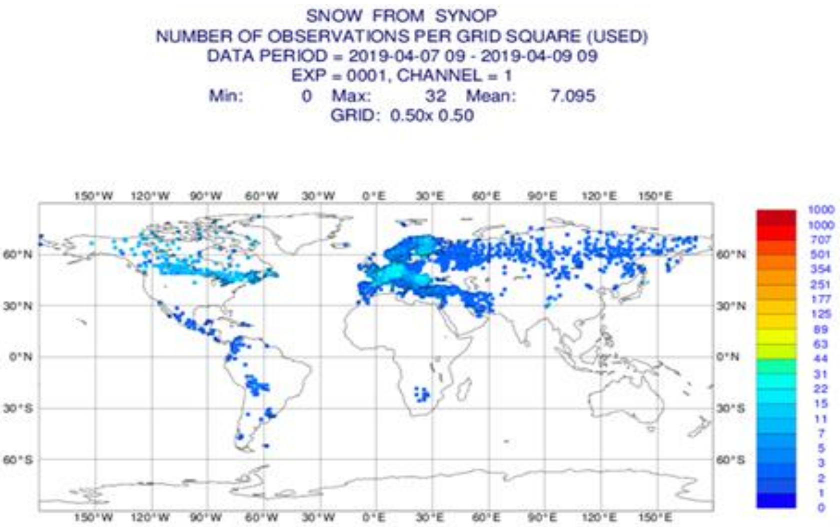

Figure 3.4.3 Number of snow observations (SYNOP) available for use in the ECMWF data assimilation

The HARMOSNOW COST Action has collected information on national practices of snow measurements, and produced an availability inventory of in Situ Snow Observation Data in Europe (European Snow booklet). This inventory covers the region shown in Figure 3.4.4 below. The number of European stations reporting snow depth data via GTS are shown in the right panel of Figure 3.4.4. The number of stations vary strongly among countries. For Iceland, Kosovo and the Netherlands, no snow data are available from GTS. For Denmark and Norway, the density of stations reporting to GTS is similar to the number of national stations. The left panel of figure 3.4.4 shows the national measurements of snow depth as reported in the European Snow Booklet (Haberkorn, 2019). The figure illustrates that quite a lot of national snow depth observations exist which are not made available from GTS.

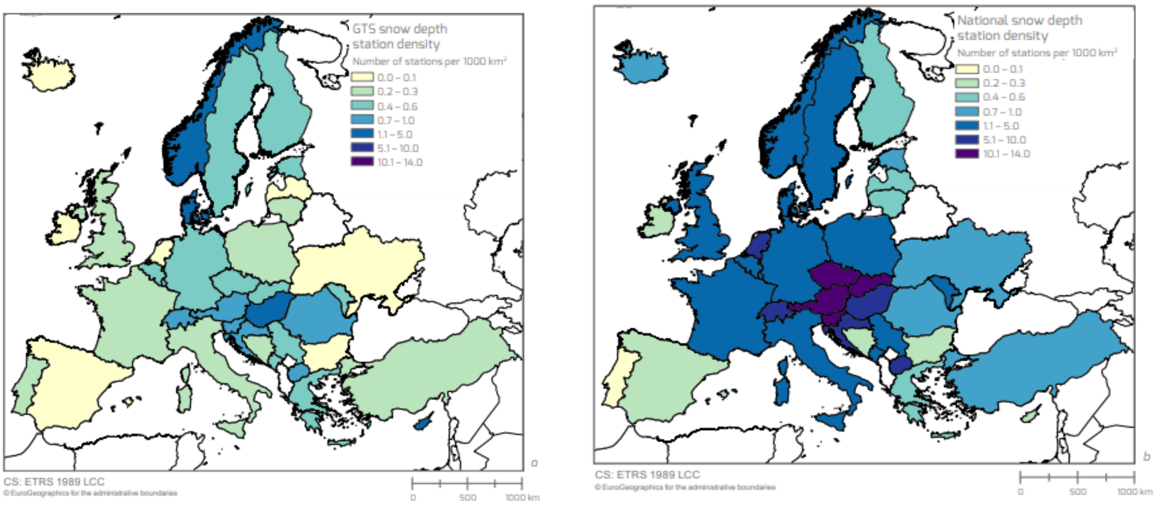

Figure 3.4.4 Maps showing the density of national snow depth stations (right) and the density of snow depth stations available from GTS (left). Figure is adapted from the European Snow Booklet (Haberkorn 2019).

The National Snow and Ice Data Center (http://nsidc.org/data) stores data from NASA, e.g. from Operation IceBridge, but also NOAA data, e.g. the World Glacier Inventory based on aerial photographs. Most of these data are not in situ data.

Operation IceBridge is a NASA project running from 2009 until 2019 with the primary goal to monitor changes in polar ice in the Arctic and Antarctic from aeroplane after the decommissioning of ICESat (October 2009) and the launch of its successor, ICESat-2, in September 2019. IceBridge aircraft carry a suite of specialized science instruments. Among these is the Airborne Topographic Mapper, a laser that measures the surface elevation of the ice. There is also a gravimeter onboard, an instrument capable of measuring the shape of cavities in the ice. There are numerous other pieces of equipment on board, including the Land, Vegetation and Ice Sensor, the Multichannel Coherent Radar Depth Sounder, a Snow Radar, a Ku-Band Radar Altimeter, a magnetometer and the Digital Mapping System.

Mapping Instruments

Digital Mapping System (DMS) - The Digital Mapping System (DMS), created by researchers at NASA's Ames Research Center, is an airborne digital imaging system that is used to detect openings in sea ice and build high-resolution maps of polar ice. The DMS instrument is a downward-facing digital camera that captures multiple individual frames that are combined into image mosaics using computer software.

Gravimeter - Operation IceBridge also uses a gravity-measuring instrument known as a gravimeter. This instrument measures the strength of gravitational fields beneath the aircraft, which researchers can use to determine the shape of water cavities beneath floating ice shelves. Because water is less dense than rock, areas of floating ice show weaker gravitational fields than areas with rock underneath.

Magnetometer - The NASA P-3 Orion carries a magnetometer that can be used to gather data on the properties of sub-ice rock. Density and magnetic properties can be used to infer bedrock type, which is helpful for determining sub-ice basal conditions.

Greenland

The Greenland Climate Network (GC-Net) consists of 18 Automated Weather Stations (AWS) distributed over the Greenland ice sheet. The GC-Net aims at monitoring climatological and glaciological parameters at various locations on the ice sheet. The measured parameters include accumulation rate and snowpack conductive heat flux, surface energy and mass balance. Currently, the GC-NET archive contains more than 50 station years of measurements. These data have been quality controlled and calibrated.

The Programme for Monitoring of the Greenland Ice Sheet (PROMICE) (www.promice.dk) is operated by GEUS (Geological Survey of Denmark and Greenland). 25 Automatic Weather Stations exist in Greenland, some of these supported by additional projects (e.g. the Greenland Analogue Project (GAP) and the Greenland Climate Research Centre (GCRC).

3.4.2. Data with restricted availability

Russian snow data

The HARMOSNOW also collected information on Russian snow measurements. The Federal Service for Hydrometeorology and Environmental Monitoring under the Ministry of Natural Resources and Ecology of the Russian Federation (ROSHYDROMET), located in Moscow, is responsible for national operational snow observations in Russia. In addition, 23 territorial administrations and 6 other research institutes coordinate additional snow observations in individual regions. Metadata is not available. A total of 1279 Russian stations deliver snow depth manually and 830 Russian stations provide snow water equivalent. Snow depth data for 24-hour periods is available at: http://aisori.meteo.ru/ClimateE Snow depth values are available for the period 1948 to 2017.

Russian snow observations may also be obtained from the World Data Centre in Obninsk (www.meteo.ru). Personal communication will be required to get access to the data.

The INTAROS research carries out fieldwork activities in the Arctic and also compiles, quality controls and makes available datasets collected by earlier campaigns, to provide access to datasets previously only available in closed communities.

The INTAROS data catalogue https://catalog-intaros.nersc.no/ contains descriptions of data collected or compiled by partners during the project. If a dataset is made available online, the link is also provided. By the end of April 2019, the data catalogue contained 42 datasets (mainly collected by project partners), but it will keep growing as the project progresses.

4. Gap analysis

4.1. Meteorology

4.1.1. Observations do not exist

The main gap in the Arctic is the lack of data and especially the lack of spatial coverage. For land meteorological observations the coordination has been weak and many independent and somehow overlapping archives exist. Many observations are only available through request, and documentation and quality control may be sparse and difficult to assess.

The lack of data and the existence of very data-sparse regions is the main gap with respect to Arctic in situ data. There are large gaps over the Arctic oceans and the inner Arctic. In particular, there is a lack of observations over northern and eastern Greenland, east of Svalbard and in particular east of the Lena delta, around the Hudson Bay, over the Labrador Sea as well as over large parts of the coastal areas of Nunavut and Yukon. In general, the SYNOP coverage over Greenland is not good.

Upper-air measurements are also very sparse in the Arctic. There are very few observations north of 70°N, no radiosonde observations at all in the inner Arctic, only 5 on Greenland and virtually none over the entire Canadian Arctic. There is also no upper air information available over the oceans as the ocean weather ships have closed down due to the high cost. A concerted action to revive one or several of them would be worth considering. For projects doing data assimilation, a limited amount of missing data is not a problem but lack of spatial coverage is. It would also be beneficial if more variables were measured.

Further, the ongoing rapid climate change creates new “hotspots” with large changes. As sea ice disappears, regions with enhanced variability may evolve in areas previously covered with ice. There is no reason to expect that these new hotspots will be found where we already have stations. Candidate regions for this kind of changes are the northern part of Svalbard and the Greenlandic east coast, both of which have very few recent observations.

4.1.2. Observations exist but data not freely available

There are many meteorological observations performed within national or community networks. These are not made freely available – most often due to lack of resources – but can most often be made available upon request. The documentation of such data is often limited.

Data obtained within research projects are in principle freely available. However, these are most often only available upon request as funding seldom covers costs related to management and publication of the data.

In particular, there seems to be a wealth of data from Soviet archives. It would be worthwhile to increase the efforts to retrieve them.

4.1.3. Technology gaps

For C3S it is important to have homogeneous time series, and it is a limitation if instrumentation and methodology change over time without proper documentation.

Comprehensive metadata as well as documentation is very important for quality assessment and proper use of meteorological data. The documentation and technical requirements are generally better if data are retrieved from a formal network with some internal guidelines. However, data from different sources are often used for the Arctic in order to gain as much data as possible.

- Data are retrieved from a large number of instruments using different technologies.
- Instrumentation may be sensitive to coverage of snow and ice and provide erroneous data. This is not easily detected. Snow covering traditional in situ temperature sensors is a major error source for Arctic temperature observations.

4.1.4. Sustainability gaps

It is very expensive to maintain Arctic stations. Some of the stations are only visited every 2nd or 3rd year, and there is a risk that all data from these stations will be unavailable for a longer period of time. Such stations are often in regions with poor coverage. There is also a fear that stations in data sparse regions will be closed because they are too expensive to maintain.

The sustainability problem has recently been surveyed and documented by the Copernicus In Situ Coordination Component (Buch et al, 2019) showing fairly good sustainability of European meteorological in situ observations, but the present analysis displays concern for the Arctic observing system.

4.2. Atmospheric composition

4.2.1. Observations do not exist

For verification with in situ data, missing data in space is inevitable. The CAMS service is interested in all sort of atmospheric composition observations both from in situ observations and ground-based remote sensing. CAMS is also very interested in observations of wet and dry deposition fluxes and precipitation chemistry as well as in dry-deposition velocities for a wide range of atmospheric constituents. Observations of pollen are also of interest for CAMS. In order to follow pollution episodes, track variability and to continuously monitor the performance of CAMS (including yearly system upgrades) it is necessary to have uninterrupted measurement records.

Biases or excessive noise in data makes it difficult to achieve goals in validation of the CAMS products over the Arctic. The limited availability of near-real time data is often more limiting than the data quality.

The current number of stations (IASOA) is limited given the large area concerned. For an adequate assessment of the transport of pollution to the Arctic during individual events, a larger number of ground stations would be helpful.

The longer the data record, the more valuable. In particular for reanalysis evaluation, long time records are needed, ideally for the full period of the reanalysis. These often do not exist.

4.2.2. Observations exists but data not freely available

Surface ozone is measured at Pallas and the data is used for reanalysis validation. It would benefit CAMS-84 if the data were available in NRT. The data owner is FMI.

4.2.3. Technology gaps

There are in general very few atmospheric stations in operation in the Arctic. Adverse weather conditions, cold temperatures, especially in winter, and prevailing darkness are among the many challenges of making measurements in the Arctic atmosphere. Logistical problems and challenges in designing instruments that can provide accurate measurements are other factors that limit the data.

It is traditionally a challenge to secure stability of data sensors in the Arctic due to the severe weather conditions and the remoteness of stations that limits the possibility for service visits. At the Villum Research Station in North Greenland, this issue has been accommodated by running two ozone monitors in parallel, so data are available from one monitor in case the other monitor breaks down. It is a costly solution, but this has dramatically improved the stability of data delivery.

Automated data control is difficult, and poor data quality can impact on verification results. CAMS reports experience of suffering from data biases, excessively noisy data, incorrect metadata and data formatting, and that there is still a gap in quality control to be addressed.

4.2.4. Sustainability gaps

The sustainability problem has recently been surveyed and documented by the Copernicus In Situ Coordination Component (Buch et al, 2019), showing severe sustainability problems European atmospheric composition in situ observations and this results is believed to be more pronounced in the Arctic.

4.3. Ocean

For CMEMS applications, there are two different data gaps:

- Observations needed do not exist. This kind of gap can be roughly identified by comparing the requirements and spatiotemporal distribution of the observations.
- Observations exist but:
    - do not fit with CMEMS purposes. As CMEMS is an operational program, most of the applications have a strict requirement for timeliness. For example, near real time forecast and validation need observations in near real time; interim reanalysis needs observations in interim scale, i.e., 1-12 months before present time. Observations for CMEMS use will also need to reach certain quality standards. This can be especially true for calibrating satellite retrieval algorithms in TACs which may need observations with very high quality.

- Data are not freely available. This can be caused by different reasons, e.g. data policy, research publication, economic benefit, technological confidentiality and even political issues.

Besides the data gaps, lack of in situ observations can be caused by technological gaps and sustainability gaps. The former relates to the technology capacity in providing operational and cost-effective monitoring for a given parameter, while the latter is determined by economic, policy, organizational and infrastructure-related issues in support for maintaining the monitoring activities.

In the following subsections, we will analyse the data -, technology - and sustainability gaps in Arctic Ocean in situ observations for CMEMS.

4.3.1 Data gaps and adequacy analysis

By evaluating data availability in section 3.3 and data requirements in section 2.3, major marine data gaps for CMEMS applications are identified as following:

Gaps due to restricted access

- Waters of the Russian Arctic: all marine data in this area is restricted for Copernicus Access

Gaps due to lack of monitoring and/or cost-effective monitoring technology

- Ice-covered waters especially in central Arctic: all marine data in the Arctic ice-covered waters are extremely sparse, only a few ITP profilers are providing operational observations. Gliders can also provide some under-ice measurements but are not yet in operational use. Cost-benefit technology such as under-ice Argo should be developed.
- The seasonal marginal ice zone: it is growing broader and none of the platforms we have today can survive there. This will be an increasing problem as the ice retreats. APB observations should be encouraged.
- Sea ice: existing datasets on sea ice (e.g. from national data centres and research projects) need to be collected and further consolidated for Copernicus use. IMB buoys are very sparse. On the other hand, the data are point measurements hence may not be suitable for validating and calibrating satellite algorithms. Near surface radiometer at weather stations can give accurate IST for satellite algorithm validation. But only a few stations exist. For the snow and ice data, air-born snow radar has played a big role in monitoring.
- Biogeochemical data: number of observed profiles from the big data centres is 1000-2000 per year for nutrient and chl-a in the Nordic Seas. However, the amount of data from non-Nordic waters is just one third to one tenth of that from Nordic waters. Some of them have restrictions on access and some of them are only released some years later. Hence these data are not fit for the purpose of CMEMS applications, especially in producing NRT forecasts and interim products.
- Wave measurements: wave data are sparse in the Arctic and are not fit for the purpose of validation of satellite products and modelling ice-wave interaction. Existing wave data need to be collated centrally and made available.

Gaps due to lack of good data management and cost-effective data transmission

- Profile data is not fit-for-purpose of NRT/interim applications: a large part of this data can only be accessed with significant time lag (years) after collection. These data do not fit the purpose of CMEMS NRT/interim applications.
- There is a lack of data collection for almost all parameters. Lots of research data and national expedition data have not been collected in Copernicus databases.

Gaps due to lack of optimal sampling strategies

- IABP ice drifting buoys are not covering the Arctic homogeneously. Satellite products provide better spatial coverage but poorer temporal coverage (data is less trustworthy in summer).
- Ice thickness measurements (IMB, BGEP) are mostly in the American Basin, the IAOOS drifting systems on the European side are not publicly available.
- Research cruises from fishery monitoring (e.g. ICES data) tend to follow fish cohorts so the sampling is seasonally biased.

In addition to above major gaps for CMEMS applications, more details on the ocean data gap analysis are given below:

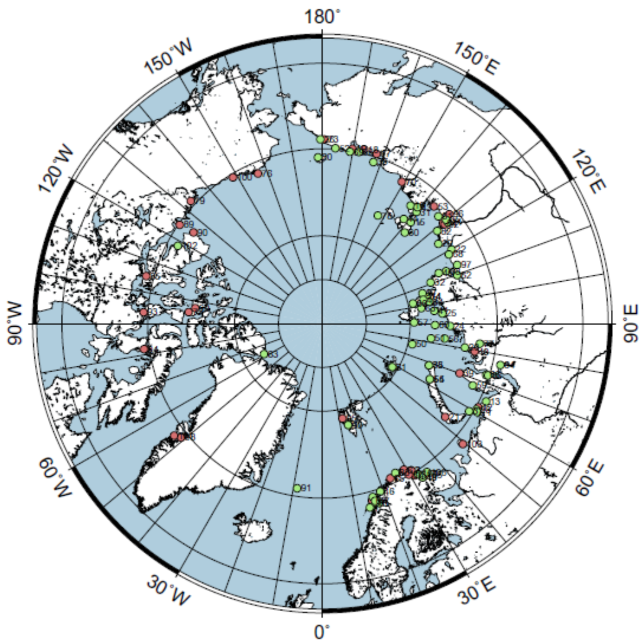

Figure 4.3.1. All 106 PSMSL tide gauges above 68° N. Green dots mark the 69 gauges with at least 5 years of data and trends within ~2 cm/year, while red dots mark rejected gauges (Source: Svendsen et al., 2015).

Sea level: currently there are a minimum of about 60 tide gauge stations operated in the Arctic coast, including 13 Russian stations. More stations with historical time series can be found in the PSMSL database (Figure 4.3.1). However only a part of the Norwegian data is included in CMEMS and only monthly data are used by ARC MFC. For SL TAC, monthly mean sea level data from 10 GLOSS stations and some PSMSL stations are used. Existing observations are adequate for CMEMS purposes. The priority is to improve data sharing, standardized format and quality control.

SST: in SST TAC, in situ SST is mainly used for calibrating and validating the satellite products. Currently only SST from about 132 surface drifters and 80 Argo profilers are used. There will be more high-quality data added from EUMETSAT/Copernicus Trusted projects. There are no significant data gaps identified for this purpose. However, many other in-situ SST datasets have been identified from the Ferrybox and moorings. For ARC MFC, regional error statistics of forecast and reanalysis are becoming important. Hence more SST data are need for validation. For data assimilation, satellite products play a major role.

IST: accurate IST can be measured by using near surface radiometers. They are needed for calibrating and validating satellite IST products. IST in situ data are sparse but there are more data existing in non-EU countries. These data can be collected and used by Copernicus.

T/S profiles: there is a good coverage of T/S profiles in ice-free waters. The data have been well collected and quality controlled and openly accessible from WOD, UDASH, iAOS, SeaDataNet, ICES and EMODnet. The estimated amount of non-Russian profiles is >20000 per year, i.e., 54 profiles per day. More than 90% of these data are open and free. This meets ARC MFC requirements of 50 profiles per day. However, data availability is not even: few data are accessible from the Russian Arctic, and few data exist in ice-covered waters. Under-ice T/S profiles are mainly observed with ITP profilers which is currently providing about 400 profiles per year (down to 500-800 m depth), as collected in CMEMS INSTAC. More ITP profiler data are needed. Regular monitoring cruise data in Greenlandic and Icelandic waters are restricted for access. In addition, there exist major data gaps for CMEMS NRT/Interim applications due to delayed accessibility of the data: the delivery time of CTD data in many cases is simply too long (1-2 years). The delayed data delivery prevents data usage for NRT and interim assimilation. One of the priorities for T/S observations is to release ship CTD data in near real time to meet the requirement for NRT and interim assimilation.

Snow and Ice over sea: the snow depth and sea ice thickness (SIT) and ice temperature can be measured by using Ice Mass Balance (IMB) buoys. An IMB array has been operated in central Artic with interruptions since 2012. However, there are only 3 stations in operation. The data gap in Copernicus can be partly filled by improved data collection but the gap is still large and caused by no data. It is also noted that, when validating satellite products, point observations still have their weaknesses in the representativeness.

Waves: currently there are about 30-40 wave buoys operating in the north of 60 N. Most of the data are found in Norwegian waters, measured at oil platforms. These data have restricted access. On the other hand, due to existence of ice, requirements on in situ wave data are higher in the Arctic than other seas for validation of satellite wave products and forecasting the wave conditions. Significant gaps in in situ wave observations exist due to both lack of data and lack of data sharing by the private sector.

Currents: a major use of current observations is for validating and calibrating the ocean models and products. Although there is a lack of current data in under-ice regions in the Arctic, there are many current profile and surface current data which have not been collected for Copernicus use. Some of these have been identified in section 3.3. HF radar can be an effective tool to fill the gaps in surface offshore currents monitoring.

Biogeochemical profiles (nutrients, oxygen, chl-a, backscattering coefficient, carbon)
Biogeochemical (BGC) in situ data in the Arctic are less than 10% of the T/S profiles. Oxygen is the most closely associated with T/S profiles. Other biogeochemical parameters are much less frequently measured then oxygen. Moorings may have oxygen sensors loaded but rarely other BGC variables. BioArgo floats can measure oxygen, chl-a, nitrate, suspended particles, downward irradiance and pH but they do not necessarily measure all the parameters. For BioArgo operated in the Arctic region, all of them have oxygen sensors but only 2-3 of them measure chl-a. Most of the BGC profiles are measured by research vessels while the data are assembled and disseminated in a late stage via ICES, SeaDataNet, WOD and GLODAPv2. This counts for about 4000 oxygen profiles, 2000 chl-a profiles and 500-1000 nutrient profiles per year. The oxygen data spread in all Arctic ice-free waters except for the Russian side. Chl-a and nutrients are mainly measured in the Nordic Seas. More data are available from iAOS but may have restricted access. In summary, even for ice-free waters, significant data gaps exist for BGC parameters, mainly due to lack of data but also due to restrictions on accessing the existing data.

4.3.3 Technology gaps

The data gaps identified in 4.3.2 are partly caused by technical gaps, i.e. there is a lack of cost-effective technology to make qualified in situ observations. For example, there are severe limitations in measurements over the seasonal ice zone and ice-covered waters. In the following we pinpoint some of technology gaps in the Arctic Ocean monitoring. Some of the issues mentioned below may not be due to the technology itself but the data gaps are caused by a specific platform:

Technologies:

- More cost-effective technology for under-ice water and MIZ monitoring, either to reduce the cost or improve the efficiency of existing technology such as ITP and gliders, or develop new technologies such as under-ice Argo, cost-effective wave monitoring instruments. This will significantly increase the amount of under-ice water measurements.
- More cost-effective BioArgo: currently a BioArgo float is 5 times more expensive than a normal Argo float. Reduced cost and improved efficiency and more parameters can be a significant contribution to BGC measurements.
- Improved sea mammals tagged monitoring: they are autonomous and thrive specifically in the seasonal ice zone. The data needs drift correction and is mostly usable in delayed-mode.
- Matured acoustic tomography technology: acoustic tomography data works in the seasonal ice zone but its use for assimilation is not yet common.
- Delivery time of cruise data: fast delivery of the cruise data is needed to fit for the purpose of CMEMS NRT/interim applications.
- More advanced data assimilation technology: this will help to use more observations such as Ferrybox data and BGC observations.

- More cost-effective technology for snow depth, ice thickness and IST monitoring: current only IMB buoys and near surface radiometers are available. The number of stations is also very much limited.
- Better technology for ice surface drifter buoys measuring IST: the ice surface drifters provide ice surface temperature. However, their quality is too low to be used for IST Cal/Val.

Data management:

- Currently there is a lack of data collection and integration for Copernicus usage. As shown in s3.3, many datasets exist but have not been collected centrally for Copernicus use.
- SST TAC: In terms of platforms, more metadata are required to better select or filter out the platforms to be used with confidence for validation. The current CMEMS NetCDF data files contain very few metadata on their associated buoy or float, while buoy or float metadata are assembled and available from JCOMMOPS web site (through an API). However, accessing both sources in near real time is hardly practicable nor very integrated.
- There would preferably be one single daily file per type of in situ source (drifters, Argo, moorings etc) rather than individual buoy files, though there are different practices among the users
- Selection of surface measurements, down to 5m depth: the proper selection of surface data should be especially investigated for Argo floats (surface profiles are often split in two or different profiles, pump off/on cases shall also be addressed)
- Better depth reporting in particular when the pressure is missing (using for instance buoy metadata) or sometimes several depths reported
- Clear identification of buoys located in sea ice areas (based on some ancillary data?) and clear indication (flag) of buoys on sea ice (and therefore reporting sea ice temperature instead of sea surface temperature). This requires dedicated data screening procedures.
- Improved quality control
  - Usage of blacklists (based on a monitoring of each buoy with respect to external sources such as satellite SST): some blacklists are already maintained by Meteo-France or UK Met Office
  - Specific checks over sea-ice areas and arctic areas (given the very high temperature anomalies in these areas over the last years)
  - Identification of “good” Argo surface measurements (as the temperature sensor may actually be off water)
  - A specific check (at least for drifting buoys) based on SST and position history. Anomalous values of SST or position have been observed in the past on drifters (e.g. position becomes lat=0 and long=0). They can be flagged out by simple tests.
  - Buoy position measurement timestamp and data measurement timestamp or a flag if they are too distant.

4.3.4 Sustainability gaps

The data gaps identified in 4.3.2 are also partly due to lack of sustainability, due to changes in conditions of resources, organizations, policies and/or political decisions. These are identified as sustainability gaps. The sustainability is relatively higher if the monitoring component is mandatory in the national or international level such as tide gauge stations, Argo floats and national environment monitoring cruises. Project-based monitoring has less sustainability.

The sustainability problem has recently been surveyed and documented by the Copernicus In Situ Coordination Component (Buch et al, 2019). The survey showed severe sustainability problems for

European ocean in situ observations in general and the results are believed to be more pronounced in the Arctic area.

4.4. Cryosphere

4.4.1. Observations do not exist

The most important observation missing:

- Too few SYNOP stations e.g. over Greenland to measure snow depth and other snow properties
- A larger spatial extent of firn cores from the Greenland Ice Sheet would be valuable for validation of satellite products
- Russian Arctic underrepresented in terms of airborne observation programs.
- More weather stations should be equipped with a GPS; this would be valuable for validation purposes.
- Additional observation data on Greenland outlet glaciers and in the ice sheet interior are needed for better validation

Much of the sea ice data (Operation Ice Bridge, aircraft missions and ice buoys) has a strong bias towards multiyear ice. In the future, ship observations from, for example, Ice Watch Assist may be of some use, but these are only covering the area in the periphery of the ice pack. There are difficulties in validating intermediate ice concentrations, in particular, as the sea ice can be very dynamic in the Marinal Ice Zone (MIZ). This causes a high uncertainty in the validation at the MIZ. Sea ice data along the MIZ and first year ice would be valuable for validating the sea ice products along the MIZ (e.g. from airborne surveys or AUVs)

4.4.2. Observations exists but data not freely available

A large number of national snow observations exist which are not made available via GTS.

There is an expectation in C3S projects that many data (e.g. Russian) are not open and available. This is, in particular, also the case for data from Soviet times.

4.4.3. Technology gaps

Snow and ice cover of instruments may give erroneous results

4.4.4. Sustainability gaps

Space agencies are a strong driver for airborne campaigns with the aim of collecting validation data for remote sensing products.

The closure of Operation IceBridge will leave a gap in the delivery of in situ observations for validation of satellite products. IceBridge performs Arctic flight campaigns and collects data of annual change in thickness of sea-ice, glaciers and ice sheets. A continuation of Operation IceBridge would be very valuable, and it is recommended that EU should contribute to future flight campaigns.

5. Discussion and conclusions

The Copernicus Services and Space component requirements for in situ data in the Arctic Region have been collected and analysed. The analysis shows that although the monitoring of the Arctic in the future will rely heavily on satellite observations, it is mandatory for Copernicus to have timely access to a broad suite of in situ observation of sufficient quality and resolution in time and space. The Copernicus community has articulated clearly which variables are essential for their production line as well as their requirement for timely delivery and quality, while the resolution in time and space is still open for further clarification. The latter issue is being addressed in the Copernicus In Situ Coordination Information System (CIS[2]), established within the Copernicus In Situ Component led by EEA.

The project group has collected a thorough, although not complete, overview of existing in situ data from the Arctic:

- Data already used by the Copernicus Services and Space component
- Data freely available at various national and international data repositories but still not used in the Copernicus production line
- Data with restricted availability due to institutional and/or national data policies
- Data collected in research projects without a data management structure enabling a free data exchange. A full overview of this data category may require further work

A gap analysis has been performed by comparing the amount of existing data to the requirements for in situ data. The analysis is general in nature since a detailed gap analysis will require clear definitions of the required resolution in time and space. The analysis has identified two groups of data gaps:

1. Observations needed but do not exist. This kind of gap can be roughly identified by comparing the requirements and spatiotemporal distribution of the observations.
2. Observations exist but are not being used because
    a. Data are not freely available due to for example. data policy, lack of institutional data management structure, research publication, economic benefit, technological confidentiality and even political issues.
    b. They do not fit Copernicus purposes due to
        i. Untimely availability - most of the applications have strong time limits, e.g. near real time forecast and validation need observations in near real time; interim reanalysis needs observations in interim scale, i.e., 1-12 months before production time.
        ii. lack of sufficient metadata
        iii. Inadequate quality - observations for Copernicus must fulfil certain quality standards.

In addition to the data gaps, two important gaps have been identified:

- Gaps in technology. The harsh and remote environment puts special demands on the instrumentation for in situ observations as well as data communication in near real-time, and the existing technology and infrastructure is extremely costly. The gap in technology has up to now put limitations on the monitoring of the Arctic, so there is a need to find innovative cost-effective technological solutions for Arctic observations securing continuous NRT data flow from this harsh environment, including during wintertime.

- Gaps in Sustainability. The sustainability of observations is highlighted as very important to maintain areal coverage and long timeseries. The sustainability problem has recently been surveyed and documented by the Copernicus In Situ Coordination Component (Buch et al, 2019) showing severe sustainability problems for European in situ observations in general - particularly for atmospheric composition and ocean observations. The present analysis shows that the sustainability issue is more pronounced in the Arctic for all thematic domains, since many in situ observations rely on time limited research funds.

General conclusions

The analysis leads to the following general conclusions:

- Environmental in situ data from the Arctic are managed by national data centres, international data centres, funding agencies and individual research project, both in countries with Arctic coastline and countries with an Arctic interest.
- National observations programmes generally meet national priorities and lack international coordination.
- The purpose of using in situ observations in Copernicus ranges from calibrating and validating satellite sensors and algorithms, numerical models to assimilation into operational and re-analysis model prognoses. In addition, for climate service, consistent and long-term observations are needed to identify the trend and long-term variability of the climate.
- In situ observations are very sparse in the central Arctic.
- Due to lack of good communication facilities many data are delivered in delayed mode and are therefore inappropriate for NRT productions in particular. Other data e.g. research data are made publicly available too late to be available even for interim re-analysis purposes i.e. there is a need for internationally agreed standards for timely delivery of delayed mode data taking into account scientists right to publish.
- The Arctic environment puts high demands on robust technology and there is a clear need to pursue innovative technology development.
- Several services express that the limited amount of data is a bigger problem than the quality of data, although poor data quality in itself is problematic.
- Insufficient data management structures at data producer level constitute a big problem which negatively affects:
    - Formats of data and metadata
    - Accessibility
    - Timely delivery
    - Quality documentation
- Access to Russian data are extremely limited and calls for a dedicated action to free more critical observations in cooperation with Russian authorities.
- The given heterogeneity of the data sources implies that:
    - Automated data quality control is difficult and poor quality can consequently significantly impact verification results

- It is important that data are collected at sites which are representative of their wider area rather than their immediate surrounding

Specific thematic domain related conclusions

Meteorology

- For land meteorological observations, the coordination has been rather weak, and many independent and somehow overlapping archives exist. Many observations are only available through request, and documentation and quality control may be sparse and difficult to assess.
- Surface observations are available from SYNOP stations, ships and drifting buoys. There are however large gaps over the Arctic oceans and in the inner Arctic. In particular, there is a lack of observations over northern and eastern Greenland, east of Svalbard, east of the Lena Delta around the Hudson Bay over the Labrador Sea as well as over the coastal areas of Nunavut and Yukon.
- Upper air measurements are important for data assimilation as well as for process understanding. Currently, the spatial distribution of stations is low, especially for higher vertical levels. There are very few observations north of 70°N and no radiosonde observations in the inner Arctic.
- Upper air measurements of more variables would be useful as well.
- Aircraft data are variable in time and space. There are very few data in the inner Arctic.
- Several weather ships have closed down due to the high expenses (“C” southeast of Greenland and “M” in the Norwegian Sea).
- The GTS system is used for effective distribution of NRT meteorological data.
- Meteorological observation data was made available from the CDS with the first full launch of data end of September 2019.

Atmospheric composition

- There is a high benefit from high-accuracy in situ data in many forms but most importantly for evaluation and quality control of various CAMS production lines. For these observations it is particularly important to have accurate estimates of observation errors.
- Most useful data for verification are derived from fixed stations
- It would be beneficial if the non-NRT data would also be made available in an agreed time frame. It would be helpful if all available data can be accessed from dedicated centres without the need to approach individual organisations.

Ocean

Major gaps in oceanographic data for Copernicus applications are identified as follows:

- There are large amounts of non-European marine data, for all key variables, which have not been collected for Copernicus use, i.e. in C3S and CMEMS in situ database.
- Openly accessible data are inhomogeneous (hence gapped) in space, time and parameters:
    - The Marginal Ice Zone, the central Arctic Ocean and other ice-covered waters are particularly undersampled areas
    - There are high density data in fishing areas;

- Data in winter are sparser than in summer;
- The availability of wave, snow and ice parameters in particular, as well as biogeochemical parameter observations, is below a critical level
- Mooring data ideally need to cover the length of a full Arctic winter.

- Near real-time delivery of CTD observations from research vessels is desirable
- The organisation of in situ data used in the CMEMS production line is handled by INSTAC, which already contains huge amounts of ocean in situ data (actually more than presently used by CMEMS MfC's and TAC's). This analysis however reveals that there exist significant amounts of freely available in situ data not yet available in INSTAC, which can be due to technical problems in establishing proper machine to machine communication, or lack of sufficient metadata or that the MFC and TAC's are not yet ready to use the data or cannot use the data due to timeliness issues - a large part of research data can only be accessed years after the data were collected and are therefore not fit-for-purpose in near real time forecast and interim reanalysis applications
- T/S profiles: Data are more than sufficient for Copernicus in ice-free waters. More data are needed in the central Arctic
- SST and sea level:
  - Data are sufficient for Copernicus service
  - Existing SST data need more comprehensive metadata. SST data from sensors on ice may cause problems. Sensors can be dropped on or trapped in the ice but metadata does not indicate this. This can cause large biases. It may be beneficial to go back to old data and check if they are affected (for long-term stability of data sources).
  - Sea level data need a common quality standard and format
  - GPS measurements associated with tide gauges are of crucial importance in order to correct sea level for vertical land motion but many tide gauges are not equipped with GPS.
- Currents: Existing data are in general sufficient but are not centralized
- Sea ice:
  - Many existing datasets on ice drift, ice thickness, ice surface temperature and snow depth/temperature have not been collected in Copernicus databases
  - The number of ice drifters with GPS is too few to derive high quality ice drift products
  - The current sources of in situ Sea Ice Thickness data are airborne surveys and ice tethered buoys and moorings. Airborne surveys are essentially snapshots and buoys only monitor at a given location. In order to be useful in a C3S context, airborne surveys should have a regional coverage of significantly more than 100 km
  - There is a significant lack, spatially and temporally, of sea ice thickness data, ice drift buoys and airborne observations for the Russian Arctic
- Waves:
  - Only a few wave buoys are running operationally in the Arctic with open data access.
  - Most of the wave data do not have open access.
  - Even with good data sharing, existing wave data are not sufficient for modelling wave-ice interactions.
- Biogeochemistry
  - Data collection is scattered and lacks coordination
  - Data are too sparse, especially in non-Nordic Seas

- The technology gap is particularly problematic in the ocean domain, there is a special need to focus on developing innovative and cost-effective technology for:
  - Under-ice water monitoring, either to reduce the cost or improve the efficiency of existing technology such as ITP and gliders, or by developing new technologies.
  - Marginal ice zone monitoring, e.g. marine mammal tagged observations, acoustic tomography technology and opportunity fishing vessels.
  - Biogeochemical, snow and sea ice variables
  - Communication of near real time data transmission and delivery
- In relation to data management there is a demand for
  - More metadata to better select or filter out the platforms to be used with confidence for validation
  - Improved quality control

Cryosphere

- A lot of national snow depth observation data exist at national meteorological services which are not distributed via GTS.
- For snow, it would be valuable if more properties were measured (e.g. snow water equivalent, snow coverage, snow grain size).
- Snow data with higher resolution (especially better coverage) are needed for requested high-resolution products
- Cryosphere observations should be representative of the surrounding region as the number of observations is limited. A better coverage of AWS on the Greenland Ice Sheet is requested. Stations need to be equipped with a GPS.
- NASA airborne operations have provided valuable data which are being used to validate satellite products in C3S. The cutting of Operation IceBridge will leave a gap in the delivery of in situ data for validation of satellite products. A European airborne survey programme is suggested as a contribution to fill this gap.
- Snow measurements will, in the longer term, be made available via Climate Data Store (CDS).

6. Next step and recommendations

The analysis has underlined the need for initiating dedicated actions to improve the availability and accessibility of critical in situ data to meet Copernicus' needs. It is however important to ensure proper coordination with other key European and international initiatives, e.g.:

- The Sustaining Arctic Observing Networks (SAON) is a joint initiative of the Arctic Council and the International Arctic Science Committee that aims to strengthen multinational engagement in pan-Arctic observing. SAON has recently approve a strategy and implementation plan covering the 2018-28 period. SAON has limited resources for implementation of and observing system but is an important high-level political player with links to national authorities.
- WMO, IOC and GEO all have a strong focus on the Arctic region due to the fact that climate changes are happening faster and are more pronounced in this region.
- EU has funded several H2020 projects with a focus on the Arctic region. These projects have collected, and continue to collect, a lot of valuable information towards design of a fit-for-purpose observation system e.g. user requirements for products and services, operation of some observation system although limited in time, etc. EU has established an Arctic Cluster for internal coordination between these projects. Especially it would be important to establish links to the INTAROS project which has a major goal to establish a roadmap towards a sustained Integrated Arctic Observations System
- Relevant national authorities since the implementation of a future sustained Integrated Arctic Observing System will rely heavily on national funding. In this respect it will also be important to investigate ways to involve indigenous people in the implementation.

As a practical way forward, Copernicus is recommended to consider the feasibility of the following short-term actions that need to be implemented in a coordinated and collaborative manner.

<u>Short term</u>

- Establish formal links to intergovernmental bodies such as SAON, WMO, IOC and GEO to ensure that Copernicus requirements for a sustained and integrated observing system are articulated and taken into account
- Liaise with Horizon Europe to promote that:
    - Arctic relevant observing technology and data communication development is included in future research calls – focus could be on multipurpose and autonomous observing platforms
    - Research projects are requested to secure a free exchange of data along the FAIR principle using existing European data management infrastructures
- Continue to setup and leverage international cooperation arrangements between the EU and non-EU countries with an Arctic interest, e.g. Canada, South Korea, Japan, and the USA
- Continue to define and document Copernicus specific requirements to an Arctic in situ observing system – attention should especially be on:
    - Resolution in space and time
    - Data quality improvement
    - Metadata improvements
- Pursue innovative cost-effective technological solutions for Arctic observations securing continuous NRT data flow from this harsh environment also during wintertime

- Initiate data rescue activities composed of but not limited to the following components:
  - Continuous support of projects like the C3S inventory effort, enhanced data collection, homogenization and mining.
  - Maintain and further develop centralised dataportals for the individual thematic domains.
  - Start a task force focussing on unlocking existing data not currently available to Copernicus. The effort could include support to organisations without a proper data management structure, and support for implementation of proper data quality control procedures.
  - Improve access to Russian data sources for all thematic areas in cooperation with relevant Russian Authorities.

Long-term actions

- Work with national authorities to:
  - Secure long-term continuation of measurements and sufficient overlapping periods when old instrumentation is replaced.
  - Stress the importance of sustained funding for a fit-for-purpose Integrated Arctic Observing System.
  - Support initiatives toward real-time delivery of expedition data for example from research vessels.
  - Increase involvement of indigenous people in data collection.
- Initiate the development of a European airborne survey programme as a contribution to fill the gap from Operation IceBridge.
- Kick-start an open-source approach toward development of a simulator tool-boxes for radar altimetry over complex surfaces such as sea ice. There is a high development and computational cost due to the small-scale variability and vast range of backscatter properties of sea ice and snow. Any initiative with this aim would be beneficial for the future sea-ice thickness algorithm evolutions.

Immediate actions

The present report was discussed at a meeting between the EEA project team and representatives from CMEMS, CAMS, C3S, ESA and EUMETSAT on 4 October 2019. Three activities that could be started immediately were identified:

1. **Prepare a catalogue of on-going and former Arctic time-limited (research) observation campaigns**.

   Objective: Make relevant observations performed by research projects during the last 10-15 years readily available to Copernicus, mainly for validation and reanalysis purposes.

   Domains: Observations related to atmospheric chemistry, cryosphere, ocean, and meteorology.

   Scope: Primarily FP7 and H2020 research projects; secondly regional and national campaigns as well.

   Responsible: EEA and members of the LOT1 consortium.

Timeframe: start as soon as possible, reporting deadline: April 2020.

2. Establish a Data Rescue TASK Force

Objective: Prepare a prioritised list of organisations and institutions identified under (1) with a view to improving the availability and accessibility of observations made during R&D measurement campaigns. The focus should be on those possessing data that meet Copernicus cross-cutting requirements.

Domains: Observations related to atmospheric chemistry, cryosphere, ocean, and meteorology.

Responsible: EEA and members of the LOT1 consortium. Input and support from relevant Services, ESA and EUMETSAT is mandatory.

Timeframe: Start March 2020, reporting deadline September 2020

3. Improve access to Russian data from the Polar region

Objective: Define and implement an approach that aims at improving the cooperation between Copernicus and Russian data custodians and owners possessing essential observations from the Polar region with the purpose of making those observations readily available to Copernicus for validation and reanalysis purposes.

Domains: Observations related to atmospheric chemistry, cryosphere, ocean, meteorology, and land.

Responsible: EEA and members of the LOT1 consortium.

Timeframe: Implementation in the course of 2020.

List of Acronyms

2D	Two Dimensional
AARI	Arctic and Antarctic Research Institute
ACTRIS	Aerosols, Clouds, and Trace gases Research InfraStructure Network
ADC	Arctic Data Centre
ADCP	Acoustic Doppler Current Profiler
ADS	Arctic Data Service, Japan
AERONET	Aerosol Robotic Network
AIREPS	AIRcraft REPort
ALTIMA	Arctic Log-Term Integrated Mooring Arrey
AMAP	Arctic Monitoring and Assessment Programme
AMDAR	Aircraft Meteorological Data Relay
AOD	Aerosol Optical Depth
AOFB	Autonomous Ocean Flux Buoys
APB	Autonomous Pinniped Bathythermograph
API	Application Program Interface
APL-UW	Applied Physics Lab. in University of Washington
ARCMFC	Arctic Monitoring and Forecasting Centres
AREX	Long-term Large-scale multidisciplinary Arctic Monitoring Program
A-TWAIN	Long-term variability and trends in the Atlantic Water inflow region
AU	University of Aarhus
AWI	Alfred Wegener Institute
AWS	Automated Weather Station
BGEP	Beaufort Gyre Experiment Programme
BGC	Biogeochemical
BSRN	Baseline Reference Surface Radiation
BUFR	Binary Universal Form for the Representation of meteorological data
C3S	Copernicus Climate Change Service
Cal/Val	Calibration/Validation
CANAPE	Canada Basin Acoustic Propagation Experiment
CDM	Common Data Model
CDS	Climate Data Store
CDR/iCDR	Climate Data Record/ Interim Climate Data Record
CEAREX	Coordinated Eastern Arctic Experiment
CLIVAR	Climate Variability and Predictability
CLRTAP	Convention on Long-Range Transboundary Air Pollution
CMEMS	Copernicus Marine Environment Monitoring Service
CNRS	Centre National de la Recherche Scientifique
CNRS-INSU	The CNRS National Institute for Earth Sciences and Astronomy (INSU)
CRISTAL	Copernicus polaR Ice, Snow and Topography Altimeter
CRREL	Cold Regions Research and Engineering Laboratory
CTD	Instrument measuring conductivity and temperature as a function of depth
DA	Data Assimilation

DBO	Distributed Biological Observatory
DFO	Fisheries and Oceans Canada
DLR	German Aerospace Centre
DMI	Danish Meteorological Institute
DOC	Dissolved Organic Carbon
DOI	Digital Object Identifier
DTU	Technical University of Denmark
EBAS	EMEP database
ECMWF	European Centre for Medium-Range Weather Forecasts
EEA	European Environmental Agency
EGO	European Glider Organisation
EIONET	European Environment Information and Observation Network
EMEP	European Monitoring and Evaluation Programme
EMODnet	European Marine Observation and Data Network
ERA5	ECMWF Reanalysis 5
ESA	European Space Agency
EUMETNET	European National Meteorological Services (a network)
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellite
EVDC	ESA Atmospheric Validation Data Centre
FRAM	Frontiers in Arctic marine Monitoring
FRM	Fiducial Reference Measurements
GAW	Global Atmosphere Watch
GC-NET	Greenland Climate Network
GCOS	Global Climate Observing System
GCW	Global Cryosphere Watch
GHCNd	Global Historical Climatology Network – Daily
GLIMS	Global Land Ice Measurements from Space
GLODAP	Global Ocean Data Analysis Project
GLOSS	Global Sea Level Observing System
GOME	Global Ozone Monitoring Experiment
GPS	Global Positioning System
GRENE	Gren Network of Excellence Program
GSN	GCOS Surface Network
GTN-G	Global Terrestrial Network for Glaciers
GTN-P	Global Terrestrial Network for Permafrost
GTS	Global Telecommunication System
GUAN	German Ultrafine Aerosol Network
HALO	The High Altitude and Long Range Research Aircraft Database
HELCOM	Baltic Marine Environment Protection Commission - Helsinki Commission
HPLC	High-Performance Liquid Chromatography
HR	High Resolution
IABP	International Arctic Buoy Program
IAGOS	In-service Aircraft for a Global Observing System
iAOOS	Integrated Arctic Ocean Observing System

IASC	International Arctic Science Committee
IASOA	International Arctic Systems for Observing the Atmosphere
ICAO	International Civil Aviation Organization
ICES	International Centre for Exploring the Seas
ICOADS	International Comprehensive Ocean-Atmosphere Data Set
ICOS	Integrated Carbon Observation System
iCUPE	Integrative and Comprehensive Understanding on Polar Environments
IERP	Arctic Integrated Ecosystem Research Program
ICSU	International Council for Science
IMB	Ice Mass Balance
IMO	Intergovernmental Maritime Organization
IMR	Institute of Marine Research
INTAROS	EU H2020 project on Integrated Arctic observation system
INSTAC	In Situ Thematic Assembly Centre.
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
IOP	Inherent Optical Properties
IOPAN	Institute of Oceanology of the Polish Academy of Science
IST	Ice Surface Temperature
ISVP	Ice Surface Velocity Drifter
ITP	Ice Tethered Profilers
JCOMM	Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology
JCOMMOPS	JCOMM Operational Prediction System
L3/L4	Level 3/Level 4
LOCEAN	Laboratoire d'océanographie et du climat : expérimentations et approches numériques
MFC	Monitoring and Forecasting Centres
MIZ	Marginal Ice Zone
MMTP	Microwave Temperature Profilers
MOSAIC	The Multidisciplinary Drifting Observatory for the Study of Arctic Climate
MSFD	Marine Strategy Framework Directive
MVP	Moving Vessel Profiler
NABOS	Nansen and Amundsen Basins Observational System
NASA	National Aeronautics and Space Administration
NDACC	Network for the Detection for Stratospheric Change
NDBC	National Data Buoy Centre
NERSC	Nansen Environmental and Remote Sensing Center
NetCDF	Network Common Data Form
NIC	National Ice Center
NILU	Norwegian Institute for Air Research
NIVA	Norwegian Institute for Water Research
NMA	Norwegian Maritime Authority
NMDC	National Marine Data Centre
NOAA	National Oceanic and Atmospheric Administration

NODC	National Oceanographic Data Centre
Norut	Norut Northern Research Institute
NPI	Norwegian Polar Institute
NRT	Near Real Time
NSF	National Science Foundation
NSIDC	National Snow and Ice Data Centre
OCTAC	Ocean Colour Thematic Assembly Centre
OSE	Observing system experiment
OSSE	Observing System Simulation Experiment
OSI	Ocean Sea Ice
OSPAR	Oslo and Paris Conventions
PAG	Pacific Arctic Group
PANGAEA	A world data centre with open access
PHOTONS	PHOtométrie pour le Traitement Opérationnel de Normalisation Satellitaire
PHY	Physics
PROMICE	Programme for Monitoring of the Greenland Ice Sheet
PSMSL	Permanent Service for Mean Sea Level
QUID	Quality Information Document
QC	Quality Control
RAN	Re-Analysis
RD	Rapid Deliver
R&D	Research and Development
RIHMI-WDC	Russian Research Institute of Hydrometeorological Information -World Data Centre
RMS	Root Mean Square
ROOS	Regional Operational Oceanographic System, EuroGOOS
R/Vs	Research Vessels
SAR	Synthetic Aperture Radar
SAON	Sustainable Arctic Observational Network
SEC	Surface Elevation Change
SEDNA	Sea Ice Experiment – Dynamic Nature of the Arctic
SIC	Sea Ice Concentration
SID	Sea Ice Drift
SIMBA	Snow and Ice mass balance Arrays
SIO	Scripps Institute of Oceanography
SIOS	Svalbard Integrated Arctic Earth Observing System
SIT	Sea Ice Thickness
SLTAC	Sea Level TAC
SST	Sea Surface Temperature
TAC	Thematic Assembly Centres
TOPAZ	Towards an Operational Prediction system for North Atlantic European coastal Zones.
T/S	Water Temperature/Salinity
TSM	Total Suspended Matter
UDASH	Unified Database for Arctic and Subarctic Hydrography
UHSLC	University of Hawaii Sea Level Centre

UiB	University of Bergen
UNIS	University Centre In Svalbard
UPTempO	Upper Layer Temperature of the Arctic Ocean
VMADCP	Vessel Mounted Acoustic Doppler Current Profiler
WARM	Warming and irRadience Measurement buoy
WCRP	World Climate Research Programme
WDCGG	World Data Centre for Greenhouse Gases
WGMS	World Glacier Monitoring Service
WHOI	Woods Hole Oceanographic Institution
WIS	WMO Information System
WOUDC	World Ozone and Ultraviolet Radiation Data Centre
WMO	World Meteorological Organization
WOD	World Ocean Database
WWW	World Weather Watch

References

Behrendt, A., Sumata, H., Rabe, B., and Schauer, U.: UDASH – Unified Database for Arctic and Subarctic Hydrography, Earth Syst. Sci. Data, 10, 1119-1138, https://doi.org/10.5194/essd-10-1119-2018, 2018. https://doi.org/10.1594/PANGAEA.872931

Biogeochemical-Argo Planning Group. (2016). The scientific rationale, design and Implementation Plan for a Biogeochemical-Argo float array, 2016.

Buch Erik, Vicente Fernández, Ines Srzic, Alex Vermeulen, 2019. Sustainability Survey. Copernicus In Situ Coordination Activity report. https://insitu.copernicus.eu/library/reports/Sustainabilitysurveyupdatedreportfinal.pdf

Haberkorn, A. (Ed.), 2019. European Snow Booklet, 363 pp., doi:10.16904/envidat.59. https://envidatrepo.wsl.ch/uploads/slf/snow-booklet/european-snow-booklet.pdf

INTAROS, 2018. Report on present observing capacities and gaps: ocean and sea ice observing system, INTAROS D2.1. 2018, pp108 https://intaros.nersc.no/sites/intaros.nersc.no/files/D2.1%20final_31May2018_0.pdf

Mader et al. (2016). The European HF Radar Inventory, EuroGOOS publications (Available at http://eurogoos.eu/download/publications/EU_HFRadar_inventory.pdf)

Mercator Ocean, EuroGOOS and CMEMS Partners (2018), CMEMS requirements for the evolution of the Copernicus In Situ Component. CMEMS Internal report. Edited by Antonio Reppucci, September 2018, pp79.

Svendsen PL, O. Baltazar Andersen and AA Nielsen. (2015). Statistical selection of tide gauges for Arctic sea-level reconstruction. Advances in Space Research 55, 2305–2314

Xie, J., Bertino, L., Counillon, F., Lisæter, K. A., and Sakov, P. (2017): Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013, Ocean Sci., 13, 123-144, https://doi.org/10.5194/os-13-123-2017, 2017.