KEPLER Deliverable Report

Report on Deliverable D5.1

Deliverable name	Synthesis on the visions of the evolution of the Copernicus services			
Scheduled delivery	**month:**	18	**date:**	May 2020
Actual delivery	**month:**	18	**date:**	June 2020 revised November 2020
Report type	Internal report			
Lead author	L. Bertino Stiftelsen nansensenteret for miljø og fjernmåling NERSC			

Contributing authors

Laurent Bertino, NERSC
Thomas Lavergne, MET Norway
Michael Gauss, MET Norway
Harald Schyberg, MET Norway
Nick Hughes, MET Norway
Annette Samuelsen, NERSC
Julien Brajard, NERSC
Michael Ying, NERSC
Steffen Tietsche, ECMWF
Gilles Garric, Mercator Ocean Intl.
Corinne Derval, Mercator Ocean Intl.
Marko Scholze, ULUND
Eirik Malnes, NORCE
Carolina Gabarro, BEC, CSIC.

Co-funded by the Horizon 2020 programme of the European Union

Table 1: Acronyms most commonly used in the present report

Acronym	Definition	Link
C3S	Copernicus Climate Change Services	https://climate.copernicus.eu/
CAMS	Copernicus Atmospheric Monitoring Services	https://atmosphere.copernicus.eu/
CDS	Climate Data Store	https://cds.climate.copernicus.eu/
CICS	Copernicus In Situ Component	https://insitu.copernicus.eu/
CMEMS	Copernicus Marine Environmental Monitoring Services	http://marine.copernicus.eu/
CLMS	Copernicus Land Monitoring Services	https://land.copernicus.eu/
EMS	Copernicus Emergency Monitoring Services	https://emergency.copernicus.eu/
EMSA	European Maritime Safety Agency / Copernicus Maritime Surveillance Service	http://www.emsa.europa.eu/copernicus.html
EOV	Essential Ocean variable	http://www.goosocean.org/index.php?option=com_content&view=article&id=14&Itemid=114
ECV	Essential Climate variable	https://gcos.wmo.int/en/essential-climate-variables

Co-funded by the Horizon 2020 programme of the European Union

Acronym	Definition	Link
ESA CCI	European Space Agency Climate Change Initiative	http://cci.esa.int/
EMODNET	European Marine Observation and Data Network	http://www.emodnet.eu/
ICES	International Council for the Exploration of the Seas	http://ices.dk
SeaDataNet	SeaDataNet	https://www.seadatanet.org/
SIAge	Sea Ice Age	
SIC	Sea Ice Concentrations	
SIThickness	Sea Ice Thickness	
SIDrift	Sea Ice Drift	
SIType	Sea Ice Type	
SWH	Significant Wave Heights	

Executive Summary

The synthesis on the visions of the evolution of the Copernicus service reports on ways to improve the description of the changing Arctic Regions in all existing and planned marine Copernicus Services capability. It is based on an inventory of polar-relevant variables (including Essential Climate Variables) that will be available in 2021 from Copernicus Services and related European databases. The inventory is made at a high level without going into technical specifications of the data products (resolution, accuracy, file formats are dis-regarded). **The report attempts to draw priorities for improved completeness and internal consistency of Copernicus services for the Arctic**.

Prioritized list of problems:

- The diversity of providers (See Annexes 1 to 5, even after narrowing down to established

Co-funded by the Horizon 2020 programme of the European Union

international sources) is prone to confusion and hampers the uptake of the most recent update of a given product. Users should not be expected to deal with the complexity of the data landscape and should be guided transparently to the best available data. We recommend thus the establishment of a one-stop-shop for all Copernicus Arctic/Polar data (across all services). It could be powered for example by a DIAS cloud solution and accesses all nominal products at their sources. Such a cross-Copernicus window should allow services as

  - Dataset discovery
  - Subsetting
  - Visualisation
  - Easy handling of polar projections.
  - Cloud computing (including the "invoke" service from INSPIRE)
  - Comparisons between different products
  - Overlays with external validation data

We suggest a few ways to improve the description of the changing Polar Regions in Copernicus Services capability:

- Fill in the red cells, which are obvious data gaps
  - Adding sea ice in situ observations
  - Additional sea ice variables from satellites.
  - Adding Permafrost variables to the CLMS and C3S data servers.
  - Adding evaporation to the CLMS data server.
  - Adding river nutrient fluxes to the ocean.
  - Adding observations of avalanches to CLMS.
- Include regional seasonal predictions of Arctic biogeochemical variables to complement CMEMS, CLMS and C3S in the ocean as well as on land.
- With lower priority, include regional seasonal predictions of ocean wave variables.
- Set up a meta-browser that can harvest polar ECV data from CMEMS, CLMS, C3S, CAMS data stores and other international sources consistently, following the example of TIGGE for weather data.
- Support international intercomparison and validation activities for ocean products (such as OceanPredict), sea ice products (such as SIDFEx) and atmospheric products (such as an Arctic focus of the SPARC Reanalysis Intercomparison working group S-RIP[1])
- Ease the transfer and/or distributed access to climate data products across programmes (C3S/CMEMS, CCI/CMEMS/C3S/, CCI/C3S/CLMS, SAFs/Copernicus) to avoid duplication, or double-branding, and merge those. Both technological and ownership aspects will need to be addressed.
- For CDRs, clarify the set of requirements yielding in the various programmes. E.g. CCI will

1 https://s-rip.ees.hokudai.ac.jp/

Co-funded by the Horizon 2020 programme of the European Union

target GCOS requirements, while SAFs and C3S will target needs of the reanalyses (that may or may not be similar to those of GCOS). Formalize the process of exchanging requirements between the CDR initiatives (See KEPLER D4.2 for more details on the latter).

Terminology

Starting from a well-established cross-disciplinary inventory of climate variables, we considered Essential Climate Variables as defined by the WMO's Global Climate Observing System (GCOS, https://gcos.wmo.int/en/essential-climate-variables/ecv-factsheets). For ocean variables, we have considered the Global Ocean Observing System (GOOS) list of Essential Ocean Variables (EOVs) www.goosocean.org/eov instead of the ECV list from GCOS.

There are subtle differences in terminology between GCOS and Copernicus documents (see table below), we have hereafter adopted the Copernicus terminology:

GCOS	Copernicus	Example
Variable	Product	Sea ice
Product	Variable	sea ice concentration

We further distinguish **reprocessing** from **reanalysis** of the variables considered,

Reprocessing: A temporally and spatially consistent history of observations without the use of a numerical model (i.e., *data-driven* only). A reprocessing involves the use of Earth Observation (satellite or in situ) retrieval algorithms and possibly spatial and temporal interpolation algorithms but not data assimilation per se. This type of observational product is often referred to as satellite Climate Data Records (CDRs) (Yang et al. 2016). CDRs are expected to cover as long a time period as raw satellite observations exist, the longest is currently 40 years long (the late 1970s). They should typically be longer than 20 years. CDRs can be continuously updated (until Y-1) by Interim CDRs. CDRs require that Fundamental CDRs (FCDRs, long time-series of calibrated raw satellite observations) are available.

Reanalysis: The application of a data assimilation procedure (including various observations into a dynamical numerical model) for a long past period of time. Examples include the ECMWF ERA-Interim and ERA5 reanalysis. A reanalysis in CMEMS has a typical duration of 25 years (starting in 1993 until the year Y-1). The C3S reanalysis ERA5 lasts from 1950 until 3 months before real-time[2] (about 70 years). Satellite-based CDRs (aka reprocessings) are typically assimilated into these reanalyses.

The above distinction is motivated by the perspective of short-term forecasting and climate scenarios,

2 A temporary ERA5T product will be available until 5-days before real time.

Co-funded by the Horizon 2020 programme
of the European Union

which require the use of numerical models and data assimilation: a reanalysis and a data assimilative forecast will use similar machinery and have therefore been assembled in the CMEMS and CAMS inventory tables. However, this distinction does not apply to the CLMS which, as defined above, only provides reprocessed variables. CDRs have a value both on their own (data-driven analysis) and as input to model-based reanalyses.

Context of deliverable within Work Package

Task 5.1 has undertaken an **inventory of Arctic products and services from Copernicus** and related European programmes. The inventory has been carried out on a high level, focused on the availability of Essential Climate/Ocean Variables i**n order to point out necessary lines of new services**. The temporal scope of the inventory is that of Copernicus-2: starting from 2021 until 2028, which is the time horizon selected for **the end-to-end roadmap in Task 5.2**. We have not discussed any developments or needs that cannot realistically be fulfilled within 2028. The "Earth System perspective" and the ongoing trend towards "seamless predictions" from days to decades are part of the motivation for this inventory because such prediction systems would require access to all updated - and mutually consistent - Copernicus data in the Arctic, as a "window" across different Services, to serve either as input (for assimilation) or as validation data. The perspective of "Digital Twin" of the Arctic should also be addressed in D5.2.

Materials and Methods

The following assumptions have been made:

- We have split **reprocessed, satellite-based, observations from model-based, data assimilative, reanalyses** because they are very different in nature and serve different purposes. Data assimilative models are to date the only way to produce forecasts, would they be on short term (1-10 days) or on a longer term (monthly, seasonal to decadal).
- We have listed the inventory by anticipation in the **post-2021 perspective**. The ongoing upgrades are therefore considered as already done and all ongoing ESA CCI (CI+ Phase 1) projects are completed.
- Differences in grid resolution, data coverage (global or regional Arctic), time coverage, choice of methods or ancillary data sources and other **technical differences between services** that can be remediated within 2021 **have been ignored**. We considered that these should not influence the post-2021 vision. Some variables require high-frequency coverage while others are fit for purpose with a snapshot every ten years (for example land biomass) so the temporal resolution of the records has been hidden.

Data sources considered

Copernicus Services

Co-funded by the Horizon 2020 programme of the European Union

- The following Copernicus services are included: **CMEMS, CLMS, C3S, CAMS**.
- Copernicus Emergency (EMS) and Security services (Maritime Surveillance by EMSA, Border Surveillance by Frontex, Support to EU External Action) are not fully included because the KEPLER consortium does not have expertise in these services. A few simple links will be noted below when deemed relevant.
- No mention has been made of satellites, sensors, instruments or other technology used or upcoming, nor of the Copernicus High-Priority Candidate Missions either. These are reviewed in detail in reports from WP3 and should be included in the roadmap D5.2.

Other data Services

- We have taken the **European perspective,** i.e. reviewing initiatives from the European Commission, not national initiatives.
- Other satellites and in situ data services have been considered that:
    - Include the whole Arctic (possibly as part of a global dataset). Where available, regional Arctic products have been preferred over global products (for example, the CMEMS Arctic over the Global products).
    - Are based on a pan-European or global collaboration. National projects are not considered.
    - Are representing the European Commission and/or the European Space Agency (ESA), and/or EUMETSAT.
    - Aggregate data from different sources and across several disciplines (i.e. not the Euro-Argo ERIC for ocean variables but rather GLODAP and EMODNET, see below). The European Ice Services are not included either since they provide one variable, albeit very important to users.
    - Are meant to be sustained. An exception has been made for the ESA CCIs, which are projects limited in time but which deliver a reference data product brokered and sometimes maintained by the Copernicus Services.
    - Some ESA CCIs are still too recent to provide validated products, we have considered in their place the precursor dataset from the respective ESA DUE project: for example GlobPermafrost.
    - Address users at a similar level: that of processed, quality controlled, geophysical values, i.e. the satellite ground segments are not taken into consideration.
- The following repositories have been selected for the ocean
    - **GLODAP** (Global Ocean Data Analysis Project): *A uniformly calibrated open ocean data product of inorganic and carbon-relevant variables*. It is an international project beyond Europe but has also been supported by European projects such as CarboOcean, CarboChange among others and constitutes the ocean contribution to the International Carbon Observing System ICOS. https://www.glodap.info/

Co-funded by the Horizon 2020 programme of the European Union

- **EMODnet** (European Marine Observation and Data Network) is funded by the European Commission DG MARE. http://www.emodnet.eu/. It includes a dedicated initiative called "EMODNET ingestion" to absorb isolated data, it has developed its own metadata standards and runs an Arctic Checkpoint project.
- **SeaDataNet** (also known as SeaDataCloud) is a pan-european infrastructure for ocean & marine data management. It federates several European (and other, i.e. Russian) in situ data centers and has a long experience in developing its own standards.
- **ICES**, the International Council for the Exploration of the Sea has historically been the first initiative to aggregate international in situ data. It focuses on fisheries and marine life but provides information on a wide range of variables.

- The following have been selected for land:
  - **ArcticGRO** (Arctic Great Rivers Observatory): https://arcticgreatrivers.org/
  - The world glacier monitoring service (**WGMS**) under the auspices of: ICSU (WDS), IUGG (IACS), UNEP, UNESCO, WMO https://wgms.ch/
  - The Global Terrestrial Network (**GTN**) with its specific servers for Permafrost (GTN-P, which contains the Northern Circular Soil Carbon Database, NCSCDv2), Glaciers (GTN-G) and Hydrology (GTN-H, and its sub-branch the GTN-R for rivers, the Arctic Runoff Data Base (ARDB)).
  - FLUXNET (Fluxcom/Fluxdata) is a vast network of meteorological sensors around the globe measuring atmospheric state variables, like temperature, humidity, wind speed, rainfall, and atmospheric carbon dioxide, on a continuous basis.
- When in situ observations are involved, different databases may be more or less complete or have different quality control practices. We have not entered these topics here as the ingestion/digestion of data can change until 2021. Our focus is more targeted to "who has the mandate to serve data?" than "who is providing most/best data?".
  - Several **other pan-European databases** are relevant for the present Arctic inventory (PANGAEA, ENACT among others) but there are so many of those that they could not have been reviewed exhaustively. We have concentrated on the main "meta-portals", the most federative for which the synchronization to Copernicus services has been on the table, but portals that only distribute meta-data have been excluded.
  - Even restricting to those, there are certain types of variables that are regionally biased towards the European Arctic (typically the in situ biogeochemical data both in CMEMS and EMODNET-Chemistry). Those have been included as they are.
  - Arctic cluster projects INTAROS and Nunataryuk do not constitute databases per se but do contribute to the databases used here. Their public deliverables as of May 2020 have been taken into account.
- The issues of **data ownership**, correct acknowledgment and traceability have not been considered here: we consider the state of available data at a given point in time and **do not exhibit the direction of data fluxes** between the different services (where is the data

Co-funded by the Horizon 2020 programme of the European Union

generated? is it brokered by Copernicus?). One of our recommendations is to ease the cross-seeding and mirroring of data repositories so that users always access the newest version (See KEPLER D4.2 for more details).

- The variables are assumed to be available on a given service, but the **quality of the service** has not been considered: **compliance to the INSPIRE directive, CARE, level of FAIR-ness or other best-practice data delivery principles**, ease of access and use. These aspects are not in the scope of the KEPLER inventory but certainly have a vital importance to the users (a data service may be public but not used because of an impractical data access).

Variables included

- We distinguish **numerical variables** (regionalized 2-dimensional or 3-dimensional variables) from integrated **indicators** (time series of spatially integrated variables, maps of anomalies from a normal situation). For example, the sea ice concentrations are a variable and the total sea ice extent is an indicator derived from it. Another example is glaciers heights (or glacier boundary) against glaciers total mass loss. The IGRAC (International Groundwater Assessment Center, https://www.un-igrac.org) provides groundwater depletion from GRACE data as a tendency by basin, but not the underlying numerical variables. CMEMS and C3S have developed numerous ocean and climate indicators, which are of broad interest, but for simplicity, we have not included them here. In other words, the **present inventory concentrates on all geo-localised numerical data** that can be presented on a map, presented in geographical details to the users and plugged into **interoperable information systems and forecast models**.

- The variables considered are the **Essential Ocean/Climate Variables (ECV, EOV)** identified by WMO/GCOS because they are already pre-selected by the scientific community for the feasibility/maturity of the observing systems. We did not attempt to include other environment and geological variables monitored under - among others - the OSPAR protocol.
- Variables relevant for **operational or tactical decisions** are included that do not qualify as ECV or EOV.
- Several **ECVs/EOVs have been ignored** because they were deemed scientifically too far from the present scope of Copernicus services (not feasible as of the present phase of Copernicus). The inclusion of these variables would entail a shift in strategy of Copernicus services. This concerns the following variables:
    - Biology and Ecosystem EOVs (plankton species, marine habitat properties).
    - Land: **Lake variables** can in principle duplicate all the ocean variables (at the exception of salinity). We have selected for the sake of brevity a few emblematic lake variables and disregarded lake 3D temperature, ice thickness, etc. **Permafrost** has several variables, which would all make identical red lines in the table, we have kept the generic name “permafrost” for the sake of brevity. The Anthroposphere ECVs

Co-funded by the Horizon 2020 programme of the European Union

(Greenhouse gas fluxes and water usage) have been discarded, as well as **indicators** that are integrated in space (glaciers mass loss, groundwater reservoirs change).

- **Ocean variables that are not yet described as EOVs or ECVs have been added to the list** because of their importance for the Arctic and because they may realistically benefit from a pan-Arctic observing system within 2028. These include sea ice variables (melt ponds, ice surface temperature, ice age, sea ice albedo, snow depths) and ocean variables (ambient noise, ocean albedo, dimethylsulfate amongst others).
- A large range of environmental variables are presently not considered as “climate” variables, such as seismic data, tsunamis, although they may be influenced by climate change in some circumstances (seismic activity related to iceberg calving in terminal fjords, seismic noise caused by waves). We have not included these numerous variables, but we note that a **European Plate Observing System (EPOS)** is proposed (Atakan et al. 2015), which is **highly relevant to EMS**. There may be practical benefits from multi-purpose in situ observatories in terms of logistics, costs and data collection that we will not discuss here.
- Some variables have been considered for inclusion but not selected because they are formally derived variables from other EOVs (for example landfast ice is a special case of sea ice drift when the latter equals zero).

Definition of time scales

The time scales defined in KEPLER D1.1 are inherited from the ice services vocabulary:

- Tactical: from present to 2-3 days.
- Strategic: from one to 10 days.
- Short-term planning: from a week to a few months.
- Long-term planning: from months to over one year.

The CMEMS and CAMS forecasts fall within the tactical and strategic time scales, while C3S forecasts address planning time scales (both short and long term).

Explanation of delays / disclaimers

The Revised D5.1 has added the following topics:

- A table for CAMS and a discussion of atmospheric reanalyses in relation to the sea ice and ocean surface.
- A colour-coding indicating the general maturity of knowledge of different variables in the Annexes, with the exception of part of the CAMS catalogue, for which we did not have the time to involve the relevant experts.

Co-funded by the Horizon 2020 programme of the European Union

Comments on the inventory

Asymmetry between Land and Ocean/Atmosphere variables

There is no table of **land model reanalysis and forecasts** since the CLMS do not use data assimilative models. Furthermore, the global CLMS product portfolio is structured into several themes among them "Cryosphere" that includes only three variables (Lake Ice Extent, Snow cover Extent and Snow Water Equivalent), however, essentially all variables provided in the global CLMS are relevant because of the land areas in Arctic regions.

A table of operational/tactical observations has been added for the ocean. For land, it is assumed that all satellite data are primarily available in near-real time and a subset of them have been processed as a climate record.

Maturity of the variables

A colour code has been applied to each variable name to indicate the general level of maturity of each variable. Observations are mature if they are abundant enough that the variable can be considered as monitored. Model reanalysis/forecasts are mature if their values agree reasonably with observations.

- Green: Mature, observations generally agree between each other and are used with sufficient frequency to quality-check the models.
- Orange: Room for improvement, large gaps in data coverage, further research is needed to understand the discrepancies either between different observations or between models and observations.
- Red: Terra incognita. Little observations, if at all. Model data come without quality insurance.

For the sake of brevity, the maturity is indicative and will not be discussed in the text.

Crosses

They indicate the mere **presence of a given variable** within the service, regardless of whether or not the product meets the requirements for accuracy, resolution, frequency set forth by the EU, GCOS, ESA or other authorities. Only a judgement of the consistency of the variable with other databases has been indicated here in the form of a colour code (green, yellow and red, see below).

Cells without a cross

They indicate the absence of a variable identified as essential. The cell may however be left without red or orange warning in the following cases:

- There is so far no explicit user demand for this variable in any of the Copernicus services
- A pan-Arctic data coverage is out of reach for scientific or logistical reasons.
- Measurements are starting / planned but would not be able to provide a decade-long

Co-funded by the Horizon 2020 programme of the European Union

reprocessing by the end of Copernicus-2 (2028)
- The variable is used in the processing of related variables but not delivered to the service for any reason (scope, accuracy, data volume or omission).

Green coloured cells "As consistent as possible":

These represent variables that are consistent across Copernicus services, consistent with the corresponding ESA CCI and synchronized across services (or will be synchronized within 2021). The consistency between the services may not be perfect but we do not see in which way the consistency can be improved in a cost-efficient way: for example the CMEMS and C3S reanalysis of sea ice concentrations use different models and data assimilation algorithms, but since they assimilate the same satellite sea ice concentrations, the inconsistency between the two products is not alarming.

In some cases, a product from one service is available in another, although possibly at a poorer resolution (for example, the C3S ocean reanalysis ORA5 is brokered by CMEMS, albeit at a coarser resolution). We have kept these cases as "green" in order to distinguish from more fundamental inconsistencies.

Similarly, a few Land variables are not identical between the C3S and CLMS (FAPAR, Leaf Area Index, Albedo and Burnt Area) without finding evidence whether the discrepancy is of concern or not. In these cases, the green colours mark which product is synchronised with the ESA CCI while the other is left at the background colour.

There are green cells without a cross, meaning that although the variable is not *provided* by a specific service, it is *used consistently* within the service that provides it (i.e. C3S winds from ERA5 are used in CMEMS in 2021 but they will not be provided by CMEMS due to different data policy).

Orange coloured cells "Partial and can be improved":

These represent cases for which the variables can be found in different services but the consistency is insufficient from a user perspective (for example sea ice drift can be found in both CMEMS and OSI-SAF but the approaches may differ significantly - different algorithms, teams, etc...). The orange colour implies that the consistency can be improved by scientific efforts or by additional observations (that are expected to become available and used within 2028). Some examples follow from the "ocean Reanalysis table":

- The ocean surface winds, precipitation (including snowfall) and surface heat fluxes from C3S reanalysis (ERA-Interim and ERA5) are used in CMEMS ocean models, but while the winds are used "as is", the precipitations and heat fluxes are bias-corrected in time-consuming tuning experiments. Therefore the winds are marked as green (as consistent as possible), the precipitations are orange (would benefit from additional research).
- Sea ice drift and thickness vary considerably from one reanalysis to another (Chevallier et al. 2015, Uotila et al. 2018) and should be consolidated in future Arctic ice-ocean reanalyses.
- The same applies to snow depths on sea ice (Uotila et al. 2018).

Co-funded by the Horizon 2020 programme of the European Union

- Sea ice age will not be part of model reanalyses in 2021 although new sea ice models can provide this information. This variable should be included in the following phase.

From the “Ocean Climate Data Records” table

- Nutrient profiles are apparently available in several databases, but contain very different collections of profiles, while a large part of the data are still missing. **A clean collection of nutrient profiles** requires a large effort in data ingestion, removal of duplicates, re-association to the quality-checked physical profile and adherence to agreed metadata standards, among other tasks.
- A sea ice drift CDR is being undertaken within the EUMETSAT OSI-SAF and ESA CCI although it will not come into an update of CMEMS as of 2021.
- An algorithm exists for sea ice age from satellite, but there is no plan for a CDR at the time of writing this draft.
- Sea ice concentrations are shown under EMODNET-Physics but link erroneously to the NRT product instead of the REP product in CMEMS.

From the “Ocean Operational/tactical observations” table: identical to the Ocean Climate Data Records

From the “Land Climate data records” table:

- Snow water equivalent as obtained from passive microwave data does not have the necessary resolution (the “hillslope” scale is about 100m).
- The Land Cover variable from the global CLMS does not contain the **adequate land cover classes for the Arctic.** The classifications should be harmonized across C3S, CLMS and the ESA CCI.
- Mass loss from Greenland Ice Sheet and Arctic glaciers are not used as freshwater input to the C3S and CMEMS reanalysis.
- Surface soil moisture is available from C3S, CLMS and ESA but the quality is poor in the Arctic. The ESA CCI data should supersede the data presently available in CLMS.
- Lake ice and Land Surface Temperature from ESA CCI should replace the data presently available in CLMS and C3S.
- The following parameters from C3S should replace the data present in CLMS: snow cover extent, land cover, surface soil moisture, and surface albedo.

From the “Land reanalysis” table (Annex 5):

- The EFAS **river runoffs** are not used as input to the ocean in the C3S and CMEMS reanalyses.

Co-funded by the Horizon 2020 programme of the European Union

Red coloured cells "Low availability and missing data"

These represent variables that require urgent efforts to significantly improve the Copernicus services, mostly by efforts of data assembly, processing, formatting and standardization.

From the ocean climate data records table:

- **In situ sea ice observations** are only collected at a national level but are not harvested by the Copernicus services.
- The **coverage of in situ biogeochemical profiles** (nutrients and Chlorophyll-a) is very poor compared to the quantity of measurements that have been undertaken.
- There is to date no initiative to construct **mature and sustained data records for several sea ice variables (melt ponds fraction, albedo, snow depths on ice)**, even though there have been demonstrations of their retrieval.
- The **surface heat fluxes** (both radiative and turbulent) are red because these are not measured by satellites over sea ice (turbulent heat fluxes are not measured over the open ocean either).

The "Ocean Operational/tactical observations" table data are identical to those from the Climate Data Records, except that sea-ice leads and ridges have been flagged for their importance to ice navigation.

From the land reprocessing table

- Permafrost variables are neither available in CLMS nor C3S.
- In situ river nutrients measurements are too scarce to be useful. We are not aware of any attempt to assemble these data for the whole Arctic. Nor are they identified as ECVs.
- Evaporation (Latent and Sensible heat fluxes) from land is not available in Copernicus. In situ eddy-covariance data from towers is available from fluxdata.org.
- In situ observations from super-sites need to become available to all the entities involved in the cal-val activities of the upcoming HPCMs. This would ensure the maximal accuracy of the processing chains and their mutual consistency.

Discussion

Predictions on operational scales and climate scales:

The provision of near-real time data and **short-term tactical forecast from CMEMS to EMS** should be further developed, for example the storm surge forecast (sea level variable in Annex 3), which would involve high-resolution coastal models, presently not part of CMEMS.

Annex 1 "Ocean Reanalysis" misses several crosses in operational and climate forecasts. In particular we noted there is no product for biogeochemical EOVs in climate projections, nor for waves or icebergs. The reanalysis of land variables (ERA5T) are available either from C3S or EMS, the latter also producing 30-days forecasts of flood and fire risk variables using ECMWF forecasts.

Co-funded by the Horizon 2020 programme of the European Union

Exchanges between land, ocean and atmosphere

Five tables have indicated in their last column the cases for which a variable may contribute to another Copernicus service. These linkages are not activated as of 2021 and represent potential cross-thematic benefits, untapped as of today:

- **From atmosphere to ocean and land surface**: The sea ice/ocean and land surfaces are extremely sensitive to changes in the atmosphere, which has a very large impact on forecast quality, but CMEMS is missing an orderly protocol to **evaluate new Arctic atmospheric reanalyses when they become available**. This evaluation should be done as a comparative exercise involving state-of-the-art atmospheric reanalyses and should not only concentrate on global atmospheric reanalyses, but should take regional reanalyses as for instance CARRA, a regional Arctic reanalysis with 2.5 km horizontal resolution, into account. CARRA is produced currently under the umbrella of C3S and lead by the Norwegian Meteorological Institute with five partner institutes. (CARRA is not pan-Arctic, but an updated reanalysis with pan-Arctic coverage is on the agenda for possible production in the next phase of C3S.) First results from CARRA indicate that this very high-resolution reanalysis is able to capture Polar Lows and other extreme events much better compared for instance to ERA5. It also applies new satellite data sets for surface properties as well as additional meteorological observations from national archives not used in ERA5. This will contribute immensely to better simulations and understanding of extreme events in the sea-ice and ocean and on land. **The comparative exercise should focus on the critical topics revealed in the tables below: surface winds, surface air temperature (heat fluxes), precipitations among others.** At the very least, **the biases should be quantified.** Another urgent knowledge gap that has to be addressed by atmospheric models is the fact that large amounts of data assimilated regularly in the atmospheric reanalyses in mid-latitudes are not taken into account in Polar regions (e.g. a large proportion of microwave sounder derived near-surface profiles are not considered by many DA systems because the radiances are not known well enough). On a less urgent note, the deposition of Nitrogen and Phosphorus nutrients to the ocean is calculated in the EMEP model, but not distributed by CAMS.
- **From ocean to land**: the **river freshwater discharges** from C3S are not used by any ocean model. This would be useful to evaluate the impact of changes in the water cycle on the ocean circulation (orange colour). The same applies even more urgently to **river nutrient fluxes** and their strong impact on the Arctic ecosystem (red colour). The EMODNET-Physics portal does list river fluxes from the HYPE model and GRDB and EMODNET-Chemistry holds links to several river chemistry monitoring systems and projects, but none of the latter appears to be relevant for the Arctic.
- **From ocean to land**: the landfast ice, waves and sea level changes are important factors in **coastal erosion**, and could be used for future scenarios as well. Other issues include salt intrusion into coastal aquifers.

Co-funded by the Horizon 2020 programme of the European Union

- **From ocean to atmosphere**: some **precursors of clouds** are generated by the ocean algae. One could note the dimethylsulfate (DMS, Lunden et al. 2010), a share of which is produced in the ocean by coccolithophores. Carbonyl sulfide (sometimes called OCS or COS) is thought to be an important **stratospheric aerosol** in the Earth radiative balance (Brühl et al. 2012), which ocean cycle (photoproduction and hydrolysis, Blezinger et al. 2000) is significant and therefore a data gap in the Arctic has been previously noted (Lennartz et al. 2020).

References

Atakan, K., Bjerrum, L. W., Bungum, H., Dehls, J. F., Kaynia, A. M., Keers, H., … Yuen, M. Y. (2015). The European plate observing system and the Arctic. Arctic, 68(5), 1–7. https://doi.org/10.14430/arctic4446

Blezinger, S., Wilhelm, C., and Kesselmeier, J.: Enzymatic consumption of carbonyl sulfide (COS) by marine algae, Biogeochemistry, 48, 185–197, 2000.

Brühl, C., Lelieveld, J., Crutzen, P. J., and Tost, H.: The role of carbonyl sulphide as a source of stratospheric sulphate aerosol and its impact on climate, Atmos. Chem. Phys., 12, 1239–1253, https://doi.org/10.5194/acp-12-1239-2012, 2012.

Chevallier, M., Smith, G. C., Lemieux, J.-F., Dupont, F., Forget, G., Fujii, Y., … Wang, X. (2015). Uncertainties in the Arctic sea ice cover in state-of-the-art ocean reanalyses from the ORA-IP project. Climate Dynamics, SI, 1–30. https://doi.org/10.1007/s00382-016-2985-y

Lennartz, S. T., Marandino, C. A., von Hobe, M., Andreae, M. O., Aranami, K., Atlas, E., Berkelhammer, M., Bingemer, H., Booge, D., Cutter, G., Cortes, P., Kremser, S., Law, C. S., Marriner, A., Simó, R., Quack, B., Uher, G., Xie, H., and Xu, X.: Marine carbonyl sulfide (OCS) and carbon disulfide (CS2): a compilation of measurements in seawater and the marine boundary layer, Earth Syst. Sci. Data, 12, 591–609, https://doi.org/10.5194/essd-12-591-2020, 2020.

Lundén, J., Svensson, G., Wisthaler, A., Tjernström, M., Hansel, A. And Leck, C. (2010), The vertical distribution of atmospheric DMS in the high Arctic summer. Tellus B, 62: 160-171. doi:10.1111/j.1600-0889.2010.00458.x

Uotila, P., Goosse, H., Haines, K., Chevallier, M., Barthélemy, A., Bricaud, C., … Zhang, Z. (2018). An assessment of ten ocean reanalyses in the polar regions. Climate Dynamics, 1–38. https://doi.org/10.1007/s00382-018-4242-z

Yang, W., John, V. O., Zhao, X., Lu, H., and Knapp, K. R.: Satellite Climate Data Records: Development, Applications, and Societal Benefits, Remote Sens., 8, 331, https://doi.org/10.3390/rs8040331, 2016.

Co-funded by the Horizon 2020 programme of the European Union

Annex 1: Ocean and Sea Ice Reanalysis table

Colours in the "Variable" column indicate the level of maturity of the reanalysis variable (agreement with observations, if any). These are omitted for variables that are not provided in reanalysis.

Colours in the other columns indicate a level of consistency between different repositories and with in situ data.

Themes	Variable	CMEMS	C3S	Short term CMEMS forecast	Seasonal C3S forecast	Cross-Copernicus value
Sea State	Significant wave heights	X	X	X		EMS, Coastal erosion
	Surf. Stress		X	X		
	Spectra	X	X	X		
Physical ocean	SST	X	X	X	X	
	SSH	X		X		EMS Floods, coastal erosion
	Surface currents	X	X	X		EMS
	Subsurface Salinity	X	X	X		
	Subsurface Temperature	X	X	X		
	Subsurface currents	X	X	X		EMS
	SSS	X	X	X		
Biogeochemical Ocean	Oxygen	X		X		
	Chl profiles	X		X		
	Nutrients (NO2, NO3, NH4, PO4, Si, Fe)	X (except NO2)		X (except NO2)		
	Zooplankton	X		X		
	Phytoplankton (PHYC+PP)	X		X		
	Coeff. Attenuation (KD)	X		X		
	CFCs tracers					C3S
	pCO2, DIC, Alkalinity, pH	X				
	Nitrous Oxide (N2O)					C3S, CAMS
	Particulate Matter					
	d13C (Carbon isotope)					C3S
	DiMethyl Sulfate (DMS)					CAMS
	Carbonyl Sulfide (OCS, COS)					CAMS

Co-funded by the Horizon 2020 programme of the European Union

Themes	Variable	CMEMS	C3S	Short term CMEMS forecast	Seasonal C3S forecast	Cross-Copernicus value
Sea Ice	SIC	X	X	X	X	EMS
	SIT	X	X	X	X	EMS
	SIDrift	X	No	X		EMS
	ISTemperature		No			
	SIAge		No			
	Melt ponds					
	Sea ice / snow Albedo					CAMS
	Ice salinity					
	Leads detection					
	Pressure ridge size and distribution					
	snow depths	X				
Atmosphere (Surface)	Winds*		X		X	
	Precipitation*		X		X	
	Ocean surface Heat fluxes (total)		X		X	CAMS
Cross-disciplinary	Iceberg density					EMS

*: Note that these atmospheric variables are not part of CAMS, contrary to what may be expected.

Co-funded by the Horizon 2020 programme of the European Union

Annex 2: Ocean and Sea Ice Climate Data Records table

Ocean Observations Reprocessing. Levels of maturity for each variable are colour-coded as previously indicated.

	Variables	Sat / in situ*	CMEMS	C3S	EUMETSAT	ESA CCI	EMODNET	GLODAP	SeaDataNet	ICES
Sea State	SWH	Sat	X	X		X			X	
	Surface stress	Sat								
	Spectra	In situ	X						X	
Ocean Biogeo	Oxygen	in situ	X				X	X	X	X
	Ocean Colour	Sat	X	X		X				
	Surface chlorophyll	Sat	X	X		X				
	Chlorophyll profiles	in situ	X				X			X
	Nutrients (NO2, NO3, NH4, PO4, Si, Fe)	in situ	X				X	X		X
	CFCs tracers	in situ						X		
	pCO2, DIC, Alkalinity, pH	in situ	X				X	X		X
	Nitrous Oxide N2O	in situ						X		
	Particulate Matter	sat	X	X		X	X			
	d13C (Carbon isotope)	in situ						X		
Sea Ice	SIC	Sat	X	X	X	X	X			
	SIThickness	Sat	Ongoing	X		X				
	SIDrift	Sat	X		Ongoing	Ongoing				
	SIType	Sat		X						
	SIAge	Sat								
	IST (Ice Surface Temp)	Sat	X							
	Melt ponds	Sat								
	Albedo	Sat								
	Ice salinity	Sat								
	Snow depths	Sat								
	Leads detection	Sat								
	Pressure ridge size and distribution	Sat								
	All sea ice variables	in situ								

Co-funded by the Horizon 2020 programme of the European Union

	Variables	Sat / in situ*	CMEMS	C3S	EUMETSAT	ESA CCI	EMODNET	GLODAP	SeaDataNet	ICES
Atmosphere	Winds	Both		X	X					
	Precipitation	in situ		X						
	Radiative H fluxes	Sat			X					
	Turbulent H fluxes	Sat								
Ocean physics	SSS	Both	X			Arctic+SSS				
	T/S profiles	in situ	X				X	X	X	X
	SSH	Both	X	X		X				
	Surface currents	in situ	X			GlobCurrent				
	subsurface currents	in situ								
	SST	Sat	X	X	X	X				
Cross-disciplinary	Iceberg density	Sat	X							
	Acoustics (ambient noise)	in situ					X			X
	Ocean bathymetry	Both					X			

Co-funded by the Horizon 2020 programme of the European Union

Annex 3: Ocean and Sea Ice operational/tactical observations table

Ocean and Sea Ice Observations in near-real time. Note that CMEMS is the only related Copernicus Service so the consistency with C3S is not an issue here.

	Variables	Sat / in situ*	CMEMS	EUMETSAT	EMODNET	SeaDataNet	Cross-Copernicus value
Sea State	SWH	Sat	X			X	EMS
	Surface stress	Sat	X				
	Spectra	In situ	X			X	
Ocean Biogeo	Oxygen	in situ					
	Ocean Colour	Sat	X				
	Surface chlorophyll	Sat	X				
	Chlorophyll profiles	in situ			X		
	Nutrients (NO2,NO3,NH4, PO4, Si, Fe)	in situ					
	Particulate Matter	sat	X			X	
Sea Ice	SIC	Sat	X	X	X		EMS
	SIThickness	Sat	X				EMS
	SIDrift	Sat	X	X			EMS
	SIType	Sat	X		X		
	SIAge	Sat					
	IST (Ice Surface Temp)	Sat	X				
	Melt ponds	Sat					
	Albedo	Sat					
	Ice salinity	Sat					
	Snow depths	Sat					
	Leads detection	Sat					EMS
	Pressure ridge size and distribution	Sat					EMS
Atmosphere	Winds	Both	Sat only	X			
	Precipitation	in situ					
	Radiative H fluxes	Sat		X			
	Turbulent H fluxes	Sat					
Ocean physics	SSS	Both	X				
	T/S profiles	in situ	X		X	X	
	SSH	Both	X				EMS
	Surface currents	in situ	X				EMS
	subsurface currents	in situ					EMS
	SST	Sat	X	X			
Cross-disciplinary	Iceberg density	Sat	X				

Co-funded by the Horizon 2020 programme of the European Union

	Variables	Sat / in situ*	CMEMS	EUMETSAT	EMODNET	SeaDataNet	Cross-Copernicus value
	Iceberg Individual	Sat	X				EMS
	Acoustics (ambient noise)	in situ					

Co-funded by the Horizon 2020 programme of the European Union

Annex 4: Land Climate and Operational Data Records table

The Climate data records are also available in near-real time.

Theme	Product / ECV	Variable	Sat / in situ	C3S	CLMS	ESA CCI	GTN and Others?	Cross-Copernicus value
Hydrosphere	River	Water level	sat + in situ	*	X			
		River discharge	in situ	*			Arcticrivers.org and ARDB (GTN-R)	CMEMS & C3S
		River nutrients	in situ				Arcticrivers.org	CMEMS
	Groundwater	Groundwater storage change	sat + in situ					
	Lakes	Lake water area	sat		X	X		
		Lake water surface temperature	sat		X	X		
		Lake colour	sat		X	X		
		Lake water level	sat	X	X	X		
		Lake ice area concentrations	sat	X	X	X		
Cryosphere	Greenland Ice Sheet	Elevation change	sat	X		X		
		Area	sat	X		X		
		Velocity	sat	X		X		
		Mass balance	sat	X		X		CMEMS, C3S
	Glaciers	outline	sat	X		X		
		elevation change	sat	X		X		
		Mass balance	sat + in situ	X				CMEMS, C3S
	Permafrost	Presence of permafrost	in situ			X**	GTN-P	
	Snow	Snow water equiv.	sat		X	X		
		Snow cover extent	sat	X	X	X		
		Avalanches	sat					EMS
Biosphere	Land Cover	Classification	sat	X	X	X		
	Surface Soil			X	X	X		

Co-funded by the Horizon 2020 programme of the European Union

Theme	Product / ECV	Variable	Sat / in situ	C3S	CLMS	ESA CCI	GTN and Others?	Cross-Copernicus value
	moisture							
	Above-ground biomass		sat			X		
	Evaporation from land	Latent & Sensible heat flux	in situ				fluxdata.org	
	Fire	Burnt Area Extent	sat	X	X	X		EMS
		Active fire	sat				GWIS [3]	EMS [4] & CLMS
	FAPAR [5]		sat	X	X			
	Leaf Area Index		sat	X	X			
	Albedo	4 bands (broadband + spectral, direct + indirect)	sat	X	X			
	Land Surface Temperature		sat		X	X		
	Soil carbon		in situ				NCSCDv2	
Others	Coastal erosion		both			GlobPermafrost		
	Ground Motion		sat		Foreseen			EMS

Notes:

*: Available in the EMS (The ECMWF EFAS system) but not as part of the C3S and thus not publicly available.

**: The ESA GlobPermafrost mean average ground temperature uses both satellite data and a model, so it should in principle belong to the "reanalysis" table but we rather kept it with the rest of the land variables because there would not be any other model-based variables in that table.

3 Global Wildfire Information System (GWIS) is a Copernicus/GEO/NASA collaboration

4 Satellite data of active fires can also include gas flares at sea.

5 Fraction of absorbed photosynthetically active radiation

Co-funded by the Horizon 2020 programme of the European Union

Annex 5: Land Reanalysis table

CLMS does not appear in this table because it does not provide numerical model output.

EMS forecasts are available for a 30-days range, so not strictly seasonal range but still within "planning" time scales.

Theme	Product / ECV	Variable	C3S	EMS	C3S seasonal forecast	Cross-Copernicus value
Hydrosphere	River	surface runoff	X		X	
		River discharge		X	EMS	CMEMS, C3S
		River nutrients				CMEMS
	Groundwater	sub-surface runoff	X			
	Lakes	Lake ice temperature	X			
		Lake water bottom temperature	X			
		Lake biology				
		Lake water temperature	X			
		Lake ice thickness	X			
Cryosphere	Glaciers					CMEMS, C3S
	Permafrost	Soil temperature profile	X			
	Snow	Snow depths	X			
		Snow density	X			
		Snow albedo	X			
Biosphere	Land Cover	Classification				
	Soil moisture profile		X			
	Above-ground biomass					

Co-funded by the Horizon 2020 programme of the European Union

Theme	Product / ECV	Variable	C3S	EMS	C3S seasonal forecast	Cross-Copernicus value
	Evaporation from land	Latent & Sensible heat flux	X		X	
	Fire	Danger indices		X		
	Evaporation from vegetation		X			
	Leaf Area Index	High/low vegetation	X			
	Albedo	Total	X			
	Soil carbon					

Co-funded by the Horizon 2020 programme of the European Union

Annex 6: Atmospheric Composition table

Theme	Variable	CAMS	Cross-Copernicus value
Aerosol (reanalyses)	Black carbon, Organic carbon, Dust, Sea salt, Sulfates	x	
Aerosol radiative forcing	Aerosol-radiation radiative forcing, Aerosol-cloud radiative forcing, and radiative effect	x	CMEMS, C3S
Fire	Fire radiative Power and biomass burning emissions	x	EMS
Greenhouse gases (forecasts)	CO_2, CH_4	x	
Greenhouse gases (reanalyses)	CO_2, CH_4, N_2O	x	C3S
Greenhouse gases (fluxes)	CO_2, CH_4, N_2O	x	C3S
Radiation	Clear-sky surface solar irradiation Total-sky surface solar irradiation (Global, direct and diffuse radiation on horizontal surface, direct normal radiation.)	x	CMEMS, C3S
Reactive gases (reanalyses)	O_3, CO, NO, NO_2, PAN, HNO_3, CH_2O, SO_2*, CH_4, OH, C_5H_8, C_2H_6, C_3H_8, O3S, HO_2	x	EMS*

*: SO_2 from volcanic eruptions is relevant to EMS but not yet part of CAMS, see the EUNADICS-AV project for a state of research: http://www.eunadics.eu/

Co-funded by the Horizon 2020 programme of the European Union