D6.4 Data Management Plan

June 2023

Charalampos Karvelis (ECMWF)
Eduardo Penabad (ECMWF)
Chiara Cagnazzo (ECMWF)
Anca Brookshaw (ECMWF)
Pierre-Antoine Bretonnière (BSC)

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement Nº 101081460.

Document information

D6.4 Data Management Plan	
Grant Agreement number	101081460
Project title	Adaptation-oriented Seamless Predictions of European Climate
Project acronym	ASPECT
Project start date	1 January 2023
Project duration	48 months
Work Package	WP6
Deliverable lead	European Centre for Medium-Range Weather Forecasts (ECMWF)
Author(s)	Charalampos Karvelis (ECMWF) Eduardo Penabad (ECMWF) Chiara Cagnazzo (ECMWF) Anca Brookshaw (ECMWF) Pierre-Antoine Bretonnière (BSC)
Type of deliverable* (R, DEM, DEC, other)	R
Dissemination level** (PU, CO, CI)	PU
Date of first submission	Day Month Year
Revision n°	-
Revision date	-

Please cite this report as: Karvelis C., Penabad E., Cagnazzo C., Brookshaw A., Bretonnière PA. (2023), Data Management Plan, D6.4 of the ASPECT project.

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them.

* **R**=Document, report; **DEM**=Demonstrator, pilot, prototype; **DEC**=website, patent fillings, videos, etc.; **OTHER**=other
** **PU**=Public, **CO**=Confidential, only for members of the consortium (including the Commission Services), **CI**=Classified

Table of Contents

List of tables

List of figures

Executive Summary

ASPECT will design and generate enhanced seasonal-to-decadal global ensemble prediction, targeting the user needs for climate adaptation, together with a set of derived information, including derived metrics for skill evaluation of means, extremes (WP2), and end-user relevant outputs (WP4 and WP5); temporally merged data from predictions and projections (WP3); a catalogue of extreme events (WP3); statistically downscaled data (WP3); specific indicators (WP4); and data may include outputs from impact models and social science data (WP5).

The purpose of the Data Management Plan (DMP) is to describe the data that will be created in the framework of the project and provide the plans for archiving and sharing the generated data. The Data Management Plan is a living document that will be updated during the project and this initial version is mainly for the benefit of the project partners.

The Data Management Plan specifies the plan for creating, exploiting, sharing, and reusing the data for ASPECT research activities. Therefore, it describes (i) the data that is produced within the project; (ii) how it is shared and where it can be found; (iii) how the produced data follows the FAIR data principles; and (iv) how partners can contribute their data to the project.

In addition, the information in this DMP about data governance, like metadata structure and data format, has the intention to provide guidelines to make the outputs of the project compatible with the standards used by the Climate Data Store (CDS), wherever appropriate. This will ensure that services and relevant datasets developed by the project could be easily considered for inclusion in the Copernicus Climate Change Service (C3S) at the end of the project.

This is the initial version of the plan. Regular updates to the plan are scheduled for June 2024 (M18), December 2025 (M36) and November 2026 (M47).

About ASPECT

ASPECT aims for the set up and demonstration of a seamless climate information system with a time horizon of up to 30 years, accompanied by underpinning research and utilisation of climate information for sectoral applications.

The goal is to improve existing prediction systems and merge their outputs across timescales together with climate projections to unify seamless climate information as a standard for sectoral decision-making. The focus is on European climate information, but ASPECT also looks more widely where there is a policy interest (e.g. disaster preparedness).

ASPECT is a project coordinated by the Barcelona Supercomputing Center (BSC). The consortium consists of eight partners and three UK-associated partners representing world-class expertise covering a range of physical science, social science and IT disciplines including climate prediction, climate projection, impacts and applications, communication, user engagement and co-development, climate service design and delivery and data management.

The consortium consists of BSC, CMCC, ECMWF, MPI, RHMZ, SMHI, UKMO, University of Leeds, University of Oxford, the University of Zagreb and Raventós Codorníu.

1 Introduction

1.1 Purpose and Scope

The Data Management Plan describes the strategies and defines how data is handled throughout the lifecycle of the project, from its acquisition to archival and beyond from the creation to exploitation. ASPECT is a challenging project that includes a heterogeneous variety of datasets and types and a large data volume; therefore, it requires careful planning and management in its development and implementation. It is necessary to have a consistent and complete plan for all aspects of the management of the data to ensure that all the requirements are met.

In formulating the DMP, the goal is that data will be generated in standard formats with proper metadata, while their storage and access will be efficient. Documentation complements the set of internal metadata of the data for its better interpretation and access. This will ensure efficient data management throughout the project's development so that data is prepared for short and long-term preservation. Additionally, a relevant objective of the DMP is to contribute to making the data Findable, Accessible, Interoperable and Reusable according to the “FAIR” principles[1].

This is the first version of the DMP to be revised during the project within Task 6.4 included in WP6. The DMP is designed as a living document that will be adjusted to the needs of the project and altered as ASPECT progresses. This initial version is aimed at the project partners.

1.2 Acronyms and Abbreviations

Acronyms and abbreviations used in this document are defined in the following tables.

Institutions and organisations

Acronym	Definition
ASPECT	Adaptation-oriented Seamless Predictions of European ClimaTe
BSC	Barcelona Supercomputing Center
CMCC	Euro-Mediterranean Center on Climate Change
C3S	Copernicus Climate Change Service
ECMWF	European Centre for Medium-Range Weather Forecasts
MPI	Max-Planck Institute for Meteorology
RHMZ	Republic Hydrometeorological Service of Serbia

[1] https://www.go-fair.org/fair-principles/

SMHI	Swedish Meteorological and Hydrological Institute
UKMO	UK Met Office
WMO	World Meteorological Organization

Acronym	Definition
CDS	(Copernicus) Climate Data Store
CF conventions	Climate and Forecast Conventions
CMIP	Coupled Model Intercomparison Project
CMOR	Climate Model Output Rewriter
CORDEX	Coordinated Regional Climate Downscaling Experiment
DMP	Data Management Plan
ESGF	Earth System Grid Federation
FAIR	Findable, Accessible, Interoperable & Reusable
GRIB	GRIdded Binary or General Regularly-distributed Information in Binary form
MARS	(ECMWF) Meteorological Archive and Retrieval System
NetCDF	Network Common Data Form
PID	Persistent IDentifiers
S2D	Seasonal to Decadal
SCI	Seamless Climate Information
WP	Work Package

2 Data Summary

2.1 Introduction

ASPECT will provide data produced by different types of numerical simulations and metrics based on these model outputs. Aside from numerical simulation data, meteorological and nonmeteorological data will be collected, corresponding to information provided by the users as part of WP5 activities.

A summary of all the data to be produced in the framework of the ASPECT project is available in section 2.2. In the current version of the DMP, only the numerical simulations are included (section 2.2.1). Additional sections will be added in the future update of the DMP for the summary of all the different ASPECT data types (e.g. catalogue of extreme events, statistically downscaled data, user and social science data, etc).

Apart from the different existing data that will be re-used (e.g. model data sets from CMIP5, and CMIP6, as well as observations and reanalyses), ASPECT will produce data from new climate numerical simulations.

Data will be produced gradually throughout the project's life and will be made available to users in a timely fashion. Every year of the project, new data from numerical simulations will be generated which will be used by other WPs in combination with the existing data to produce new derived data and products, and eventually, the project will create data for the end users.

2.2 Data Produced under ASPECT project

2.2.1 Earth System Models Output Data

In ASPECT, existing or new prediction and projection model simulations will be used to produce gridded output data. In particular, the following simulations will be run:

- seasonal predictions (up to 7 months);
- extended seasonal predictions (up to 36 months);
- decadal predictions (up to 10 years);
- extended decadal predictions (up to 20 years);
- 30-year initialized outlooks (up to 30 years);
- centennial climate projections (1850 - 2100).

Simulations using similar protocols to seasonal predictions but extending over a period of several years will be referred to as 'extended seasonal predictions'. The decadal prediction will be under the DCPP experiment family[2] following the CMIP6 protocols. At a later stage, models will follow the CMIP7 experimental design. Protocols for the 30-year initialised outlooks will be defined at a later stage.

A brief description of the experiments produced by the individual centres is provided in Table 1.

Production

Model outputs will be generated by running a number of numerical weather and climate models in several different numerical simulations from the participating modelling centres. Table 1 describes the experimental design deployed in WP1.

[2] https://view.es-doc.org/?renderMethod=id&project=cmip6&id=ecdd7e96-a807-436d-a2db-bb6efb172558&version=1&client=esdoc-search

Table 1a: ASPECT Experimental Design (seasonal predictions)

Centre	Period	Real-time forecast ensemble size	Hindcast ensemble size	Model Version
ECMWF	1981 - 2016	51 members	25 members	SEAS5[3]
			51 members for Feb/May/Aug/Nov	
CMCC	1993 - 2025	60 members	40 members	Seamless System (new system, under development)
UKMO	1993 - 2016		80 members	GloSea6[4]

The length of the forecast is 7 months and the models are initialised on the 1st of the months of each month.

In the first phase of the project, the new high-resolution CMCC system will be used only for the seasonal forecasts, whilst the remaining forecast horizons (extended-seasonal, decadal and extended-decadal) will be covered in a seamless way using the current operational decadal prediction system (CMCC-CM2-SR5).

Table 1b: ASPECT Experimental Design (extended seasonal predictions)

Centre	Length of forecast and Hindcast	Initialization	Period	Real-time forecast ensemble size	Hindcast ensemble size	Model Version
ECMWF	12 months	Feb/May/ Aug/Nov 1st of the month	1981 - 2016		15 members	SEAS5[3]
ECMWF	24/28 months	May/Nov 1st of the month	1960 - 2025		20 members	SEAS5-low res[3]
						SEAS6-Prototype

[3]https://www.ecmwf.int/en/newsletter/154/meteorology/ecmwfs-new-long-range-forecasting-system-seas5

[4]https://www.metoffice.gov.uk/research/climate/seasonal-to-decadal/gpc-outlooks/user-guide/global-seasonal-forecasting-system-glosea6#:~:text=GloSea6%20stands%20for%20Met%20Office,will%20occur%20in%20two%20phases

BSC	36 months	May/Nov 1st of the month	Nov init: 1960-2021 May init: 1980-2021		20 members	EC-Earth3[5]
					15 members	EC-Earth-3-HR[5]
CMCC	24/28 months	May/Nov 1st of the month	1980 - 2025		20 members	CMCC-CM2-SR5[6]
						Seamless System (new system, under development)
UKMO	24/28 months	May/Nov 1st of the month	1980 - 2022		> 20 members	GloSea6[4] (TBC)

Table 1c: ASPECT Experimental Design (decadal predictions)

Centre	Period	Real-time forecast ensemble size	Hindcast ensemble size	Model Version
BSC	1960 - 2022	20 members	20 members	EC-Earth3[5]
BSC	1960 - 2025	30 members	30 members	EC-Earth4[7]
CMCC	1960 - 2025	30 members	20 members	CMCC-CM2-SR5[6]
				Seamless System (new system, under development)
MPI	1960 - 2019		80 members	MPI-ESM-Low res[8]

[5] https://ec-earth.org/ec-earth/ec-earth3/

[6] https://gmd.copernicus.org/articles/16/179/2023/

[7] https://ec-earth.org/ec-earth/ec-earth4/

[8] https://www.wdc-climate.de/ui/cmip6?input=CMIP6.DCPP.MPI-M.MPI-ESM1-2-LR

UKMO	1960 - 2025		40 members	GC5[9]

The length of the forecast is 10 years and the models are initialised on the 1st of November of each year.

In a later stage of the project, when the CMIP7 forcings will become available, the new high-resolution CMCC system will be used for all the forecast horizons (seasonal, extended-seasonal, decadal, and extended-decadal).

Table 1d: ASPECT Experimental Design (extended decadal predictions)

Centre	Period	Real-time forecast ensemble size	Hindcast ensemble size	Model Version
BSC	1960 - 2025 every 5 years			EC-Earth4[6]
CMCC	1960 - 2025 every 4/5 years		10 members	TBC
MPI	1960 - every 5 years		16 members	MPI-ESM-Low res[7]

The length of the forecast is 20 years and the models are initialised on the 1st of November every 4-5 years.

Table 1e: ASPECT Experimental Design (30-year initialised outlooks)

Centre	Length of forecast and Hindcast	Initialization	Period	Real-time forecast ensemble size	Hindcast ensemble size	Model Version
ECMWF	30-year initialized	Nov 1st of the month	1985, 2005 and 2025		10 - 15 members	SEAS6
BSC	30 years	Nov 1st of the month	1960 - 2025 every 5 years			EC-Earth4[6]

The length of the forecast is 30 years, and the models are initialised on the 1st of November. ECMWF will initialise the model for 3 start dates, while BSC every 5 years.

[9] https://centaur.reading.ac.uk/112173/

Table 1f: ASPECT Experimental Design (climate projections)

Centre	Length of forecast and Hindcast	Period	Real-time forecast ensemble size	Hindcast ensemble size	Model Version
BSC		1850 - 2020		15 members	EC-Earth4[6]
MPI	Historical simulations	1850 - 2015		50 members	MPI-ESM-Low res[8]
MPI	Scenario Projections	2015 - 2100		50 members	MPI-ESM-Low res[8]

The list of variables that will be produced is available in appendix III.

Repositories

ASPECT data is deposited in public repositories for sharing internally to the project. The model outputs of seasonal and extended seasonal predictions will be securely archived within the ECMWF MARS system. On the other hand, decadal and extended decadal predictions, along with 30-year initialised outlooks, will be archived at the local ESGF data nodes of the centres. In all cases, the data will strictly adhere to the archiving system's requirements. Details about the data access to these repositories can be found in the data access section of this document and in Appendix I.

Both MARS and ESGF are online archiving systems to store and distribute data. The selected archiving systems are well-known, trusted and very robust storage infrastructures for weather and climate data. Through these public repositories, data will be made openly available ensuring that the climate protocols and metadata have been followed. Additionally, this approach facilitates a smooth transition, enabling the data openly and widely accessible beyond the ASPECT project.

At the end of the project, C3S may decide to publish a few selected datasets in a dedicated CDS repository if they would consider there was a sufficiently strong user demand to make those available. The final decision regarding publication rests with the C3S management board.

A plan about which model outputs will be archived and when they are made available will be part of the future update of the DMP.

Data Formats and Conventions

In ASPECT, the model output will be encoded in NetCDF[10] and GRIB[11] format. NetCDF is a format that is freely available and commonly used in the climate modelling community (how to read a NetCDF). GRIB is the WMO binary standard format for the international exchange of meteorological gridded data (GRIB how-to examples).

[10] https://www.unidata.ucar.edu/software/netcdf/

[11] https://library.wmo.int/doc_num.php?explnum_id=11283

The selection of these formats is based on the requirements of the archiving system and the usual standards of the community. MARS requires data to be stored in GRIB format, while ESGF in NetCDF following CF[12] and CMOR[13] conventions conform to the CMIP6/7 controlled vocabulary, with filenames described in the CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV's document. For the data which will be stored in ESGF the use of the C3S encoding recommendations for initialised ensemble decadal predictions will be promoted and it is highly recommended (more information about C3S decadal encoding recommendations is provided in Appendix II).

Seasonal and extended seasonal predictions (extended seasonal in the sense that simulations follow similar protocols to seasonal simulations but run for a longer period) are encoded in GRIB, while extended-seasonal (multi-year), decadal, extended decadal and 30-year initialized outputs in NetCDF following CF and CMOR conventions.

The NetCDF files will be created using the NetCDF Classic Model in NETCDF4_CLASSIC format (i.e. compression is allowed, but not groups and compound data types), will contain a single physical variable (along with coordinate/grid variables, attributes and other metadata) from a single model and a single simulation (i.e. from a single ensemble member and a single start date). The encoding of the model outputs in GRIB format will be done following the rules and practices for long-range forecasts at ECMWF.

In addition to the mentioned formats, the seasonal and extended seasonal predictions can also be produced in NetCDF format following the CF and C3S conventions[14] and filenames[15]. The aim of this data formatting is to align with current practices at C3S seasonal forecast activity for data delivery, allowing WP6 to build on top of C3S infrastructure and tools to acquire the data. This data will be delivered to WP6 in order to be converted into GRIB and stored in MARS. For the delivery to WP6, the files will be packaged in tar format following the principle described in the NetCDF standards for C3S Seasonal forecasts. The original NetCDF files will be internal to the project and will not be published.

Tools

GRIB and NetCDF are widely used formats for encoding gridded model outputs and in general environmental data. Various open-source software packages exist that can read and generate GRIB and NetCDF datasets.

ecCodes[16] is a tool for decoding and encoding GRIB messages, while netCDF libraries are freely available for the generation of CMOR[17] and CF-compliant NetCDF files. Well-established community tools like Python xarray, and cfgrib are available allowing users to read and process both formats in a unified approach. WP6 will also investigate solutions when needed, and develop and provide tools to improve the interoperability and usage of the data (e.g. conversion of GRIB data into NetCDF).

[12] http://cfconventions.org/cf-conventions/cf-conventions.html

[13] https://cmor.llnl.gov/

[14] https://confluence.ecmwf.int/display/COPSRV/Guide+to+NetCDF+encoding+for+C3S+providers

[15] https://confluence.ecmwf.int/display/COPSRV/NetCDF+Dataset+Design+Overview

[16] https://confluence.ecmwf.int/display/ECC

[17] https://cmor.llnl.gov/

Quality Control

The model output needs to be subjected to quality control on metadata, encoding and formatting before it can be archived. The quality control process ensures that all files meet the required standards, making data interoperable and reusable.

To avoid transferring large volumes of incorrectly formatted data, quality control needs to take place close to the origin of the data. Quality control can be done by using dedicated software.

For data to be archived in ESGF, PrePARE is a free tool provided by PCMDI (Program for Climate Model Diagnosis and Intercomparison) to validate if files conform to the CMIP6 data protocol, while nctime is the checker for the description of the time axis.

The compliance with C3S convention is checked with **c3schecker.** This is a tool developed by ECMWF and used to validate compliance of seasonal and extended seasonal forecast outputs in NetCDF with the C3S encoding standards. The c3schecker can be provided by ECMWF on request.

Data that doesn't follow the requirements of the archiving systems cannot be ingested and archived (e.g. ESGF requires the NetCDF files to be CMOR compliant, and MARS requires data to be in GRIB format). In the case of data to be delivered to ECMWF, depending on the stage where the errors are identified, data providers will be contacted to correct the errors and release a new version of this data. For the data archived in ESGF, potential errors are handled by the ESGF Errata Service[18].

However, in cases of minor technical errors which do not prevent archiving the data will be published and, whenever possible, the errors will be either documented or corrected by the data producers or the publisher depending on the error. The documented errors will be part of the general documentation accessible by all the project participants (e.g. dedicated location reachable from the project's Wiki page). The scientific content of the data is not automatically checked, any errors or issues detected during the project will be documented.

Data access

The way to access the data differs depending on the archiving system. After their publication, data will be searchable based on the available metadata and then will be available for download. Data archived at ESGF (decadal and extended decadal predictions and 30-year initialised outlooks) becomes fully accessible as soon as published through ESGF nodes. Seasonal and extended seasonal forecasts will be made available for internal project use through MARS.

Data archived in ESGF can be browsed through the ESGF portals[19] (index nodes), which are an interface for users to access data that are distributed in several ESGF data centres, while ESGF offers detailed documentation on its User Support web page[20].

[18] https://es-doc.github.io/esdoc-errata-client/

[19] https://esgf.llnl.gov/nodes.html

[20] http://www.esgf.io/esgf-user-support/index.html

Data archived in MARS can be browsed through the MARS web portal or can be accessed through the ECMWF Web API. Details about the recommended interface will be provided in a future update of the DMP.

A summary view of the formats and conventions followed for each experiment type can be found in Table 2.

Table 2: Experiments, formats, and archiving systems in ASPECT

Type of experiment	Institution	Format	Archive	Access
Seasonal Predictions	CMCC	GRIB (Data delivered to ECMWF in NetCDF format, internal to the project files)	MARS	MARS request
	ECMWF	GRIB	MARS	MARS request
	UKMO	GRIB (Data delivered to ECMWF in NetCDF format, internal to the project files)	MARS	MARS request
Extended Seasonal Predictions	BSC	NetCDF	ESGF	ESGF portals
	CMCC	TBC	TBC	TBC
	ECMWF	GRIB	MARS	MARS request
	UKMO	TBC	TBC	TBC
Decadal Predictions	BSC	NetCDF	ESGF	ESGF portals
	CMCC	NetCDF	ESGF	ESGF portals
	UKMO	NetCDF	ESGF	ESGF portals
Extended Decadal Predictions	BSC	NetCDF	ESGF	ESGF portals
	MPI	NetCDF	ESGF	ESGF portals
30-year	BSC	NetCDF	ESGF	ESGF portals

initialized outlooks	ECMWF	NetCDF	ESGF	ESGF portals
Climate Projections	MPI	NetCDF	ESGF	ESGF portals

Schedule for Data Sharing

The current document describes the plan for archiving and publishing the data for use. Detailed information on the timeline of production and availability of new ASPECT data will be published on a project wiki page[21].

Documentation

As described above the data are self-documented to some extent with the provision of the metadata and standards, to make it possible to use outside the context of this project.

However, to help the users access and correctly interpret new and existing data relevant to ASPECT, detailed documentation will be provided by WP6. Documentation will contain a list of the datasets, the models and the experiments, a definition of variables and frequencies, details on where to find the data, how to access and download the data, possible issues with the data, etc. Data documentation is a continuous process throughout the project.

User Support

First point of contact for issues related to the data is provided by WP6. This DMP does not address the resolution of issues pertaining to the data content and should be tackled by the data producers.

2.3 Existing Data used by ASPECT project

Achieving ASPECT's objectives will involve accessing a large number of different existing data sources. ASPECT will make use of data from existing weather and climate model datasets (e.g. C3S, CMIP5, CMIP6, CORDEX), and observational and reanalysis datasets (e.g. GPCP, E-OBS, HadCRUT, ERA-5, C3S regional reanalyses).

Data will be available from their original repositories under the current access conditions (e.g. C3S's User Requirements Database[22]), while WP6 will provide documentation to explain where to find and how to access this data.

3 FAIR Data

In line with the EU's guidelines, ASPECT will make an effort to follow the FAIR principles for research data (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention). The model outputs policy is for data to be free and as openly accessible as possible.

[21] https://earth.bsc.es/aspect/doku.php?id=start

[22] https://climate.copernicus.eu/user-requirements-gathering-and-analysis

3.1 Making Data Findable

In order to ensure findability, ASPECT promotes the use of globally resolvable Persistent Identifiers (PID). Data will be identified by a DOI to make content easily and uniquely citable. Please note, DOI is an application of globally resolvable PID and is external to ASPECT procedures. For data that cannot be assigned a DOI, each contributing centre has to support unique and persistent identifiers. The publication in ESGF includes a DOI (for the files) and a PID (for the datasets) creation[23].

ASPECT data makes use of well-established standards and conventions, while is archived in well-known and very robust storage systems. In that way, data will be easily discoverable and findable. All the metadata and keywords required by the archiving system will be provided, while the metadata will be offered in such a way that it can be harvested and indexed.

Information about the selected repositories and metadata that will be used are provided in Section 2 of the current document.

3.2 Making Data Accessible

ASPECT data will be accessible as soon as published in ESGF and/or MARS. Both archiving systems are trusted, robust and widely used for the storage and publication of gridded outputs of numerical models. MARS is provided by ECMWF and can be used to store, explore, and retrieve meteorological data in GRIB. ESGF is an international federation for the management, dissemination, and analysis of model output contributing to CMIP. Both ESGF and MARS allow the generation of PI such as DOI.

In addition, ASPECT is committed to strengthening the European Research Area through the open sharing of its results through open-access publications and data.

A strategy for intellectual property rights (IPR) will be adopted following the guiding principles of Horizon Europe both on openly accessible research and effective exploitation of results, to make all research outputs freely available wherever possible (allowing for any pre-existing IPR incorporated into the research outputs, including commercially sensitive information from the case study super-users).

Information on how ASPECT data can be accessed as well as restrictions on use is described in Section 2 and Appendix I, and it will be appropriately documented for the project partners (and is reachable from the relevant section of this document).

Also, note that all data produced as part of ASPECT will be submitted to the archives before the end of the project.

3.3 Making Data Interoperable

Standardisation is a key factor in making the data interoperable. Interoperability of the data can be ensured by the usage of common metadata vocabularies, encoding standards, conventions and formats. ASPECT will follow well-established practices and methodologies for the generation and usage of model outputs, already used in CMIP and C3S.

[23]https://esgf-data.dkrz.de/projects/esgf-dkrz/pid

The model output will be encoded in widely used formats following well-established conventions like C3S, CF and CMOR. At the same time, the selected repositories support the archiving and distribution of data in the format described in Section 2. In that way, data outputs will be formatted to meet requirements for data sharing and preservation.

Moreover, through the CF conventions and GRIB WMO table-based format, requirements about the structure and annotation can be satisfied (e.g. identification of variables/parameters through CF standard names, etc) which serves the interoperability of the data.

Details on the encoding standards, data format and data metadata are described in section 2.

3.4 Increase Data Reuse

ASPECT data policy is open sharing and access to data. Each dataset needs a licence attached (e.g. archived data using Creative Commons or equivalent). The recommendation in ASPECT is to use Creative Commons open licence[24] making them fully reusable and hosted through freely accessed repositories in public domain to permit the widest re-use.

In addition, documentation for individual model data will be available on the ASPECT wiki page as a project report. Documentation will be based on predefined templates, which will be filled by project partners responsible for the model simulations.

Documentation will include information on: (i) where to find the simulations and data on the repositories; (ii) description of model simulations; and (iii) any other relevant aspects, such as similarities between initialised and non-initialised experiments (e.g. 30-year integrations).

Model simulations can also be used outside the framework of the ASPECT project, for example, some of the numerical simulations constitute contributions to CMIP6/7 or other, future projects.

4 Other research outputs

Products and results of scientific analysis using tools and software developed within the project are out of the scope of the data management. The ASPECT wiki page will be used until the publication of the results.

Model code is out of data management scope and the relevant WPs will be responsible for documenting and publishing code improvement.

5 Allocation of Resources

Data Management in ASPECT is included as Task T6.4 of WP6, led by ECMWF, who will be responsible for the data management activities (e.g. defining metadata and data formats,

[24] https://creativecommons.org/licenses/by/4.0/

solutions for data access and data storage, writing documentation, etc). WP6 is also responsible for FAIR data management.

As outlined in the previous sections, ASPECT data will make use of already existing services and infrastructures at the contributing centres. The selected archiving infrastructure (MARS and ESGF) can ensure long-term preservation. User support (e.g. update of the documentation) or any other data management activity will be provided for the duration of the project.

6 Data Security

Before the data will be uploaded to the aforementioned repositories (ESGF and MARS), datasets will be stored in the partner's storage systems. Every partner is responsible for ensuring that the data are stored safely and securely and in full compliance with European Union data protection laws for the EU countries or national data protection laws for the non-EU countries.

The archiving infrastructures selected as common repositories for the project are dedicated infrastructures for storing weather and climate model outputs. They are robust archives operated by experienced groups with a long history of meteorological data management.

Once the data has been uploaded to the repository, even after the completion of the project, all the responsibilities concerning data recovery and secure storage will go to the operators of the storage infrastructure. The redundancy of the data will follow the redundancy policy of the archiving infrastructure (i.e. MARS and ESGF). Finally, the availability of the data is subject to the archiving systems.

Uploading the data to the ESGF requires storage for at least 10 years, but in practice, it will be kept for longer.

Research projects data archived in MARS do not have a pre-defined retention period as the aim is to keep the data indefinitely. This gives a sufficient level of confidence in the data being retained for a long period after the end of the project, and it is worth noting such data is not subject to the "ECMWF archive policy" which implies that in the extremely rare eventuality of a need to delete research project data, this will be assessed on a case-by-case basis in agreement with the data producer.

7 Ethical Aspects

ASPECT complies with ethical principles (including the highest standards of research integrity as set out in the ALLEA European Code of Conduct for Research Integrity, as well as applicable international and national law, including the Charter of Fundamental Rights of the European Union and the European Convention on Human Rights and its Supplementary Protocols. Appropriate procedures, policies and structures are in place to foster responsible research practices, to prevent questionable research practices and research misconduct, and to handle allegations of breaches of the principles and standards in the Code of Conduct.

ASPECT activity doesn't involve any processing of personal data or previously collected personal data.

References

European Commission, Research & Innovation, Participant Portal H2020 Online Manual, http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/openaccess-data-management/data-management_en.htm

Horizon 2020 FAIR Data Management Plan (DMP) Template

ASPECT Grand Agreement, Project: 101081460 — ASPECT — HORIZON-CL5-2022-D1-02

ASPECT Description of Action (DoA) Project: 101081460 — ASPECT — HORIZON-CL5-2022-D1-02

Appendix I Data Storage Infrastructures

MARS

All the data that will be produced in GRIB format will be stored in MARS (**M**eteorological **A**rchival and **R**etrieval **S**ystem).

MARS is a reliable and efficient tool to download data with an easy monitor of the transfers being made.

MARS has the following features:

- facilitie to Archive and Retrieve environmental data
- is a 24/7 service
- batch and interactive modes are supported
- large amount of data, both in size and number of items stored
- large number of users with different requirements.

In addition, MARS has detailed and very extensive documentation and a detailed web page on how to access archive datasets.

MARS data can be accessed through the ECMWF Web API (or MARS request). This Service will allow authorised users to **retrieve** and **list** MARS data from **outside the ECMWF facilities**. Detailed documentation is available on how to download and install ECMWF Web API as well as how to use this service.

ESGF / CMIP6[25]

ESGF (**E**arth **S**ystem **G**rid **F**ederation) is an international collaboration for the software that powers most global climate change research, notably assessments by the Intergovernmental Panel on Climate Change (IPCC) and it is designed to handle large-scale data management for worldwide distribution. It is recognized as one of the leading infrastructures for the management and access of large distributed climate data volumes and among others supports the Coupled Model Intercomparison Project (CMIP).

ESGF is a Peer-to-Peer (P2P) system with the following features:

- supports current CMIP6 activities, and prepares for future assessments;

[25]Coordinating an operational data distribution network for CMIP6 data (DOI: 10.5194/gmd-2020-153) https://gmd.copernicus.org/preprints/gmd-2020-153/gmd-2020-153.pdf

- manages several petabytes of data distributed in several federated **data centres**, also called **data nodes**;
- develops data and metadata facilities for inclusion of observations and reanalysis products for CMIP6 use;
- enhances and improves current climate research infrastructure capabilities through the involvement of the software development community and adherence to sound software principles;
- decentralized databases for handling climate data at federated sites worldwide.

Please note that although is not a climate service, the data available from ESGF is of great usefulness for impact assessment scientists and other climate services users too.

The ESGF User Support web page can be found here.

ESGF has two main components, the data nodes (or data centres) and the index notes (or ESGF portals). ESGF data nodes are used to store data which is then published to the index node. In other words, the ESGF portal is an interface for users to access model data that are distributed in several data centres. Please note that the portal itself does not host any data but is only used to facilitate the search and download. The list of nodes with portals (index nodes) can be found in the link to the portals.

How to upload a file

In order to upload a file in ESGF please follow the instructions on the ESGF User Support page.

How to access data

ASPECT data stored in ESGF can be downloaded in several ways, depending on the needs of the user, it can be a manual download for a small number of data or semi- automatically for larger number of data.

The first step is to search for data of interest through the ESGF portal (index nodes). Detailed instructions on how to make a data search and download are available on the ESGF User Support Page.

The first very basic way to download a single file:after the search for data, click on “list Files” and then “HTPPServer”, authenticate and download the file via the web browser.

An alternative more advance way to download files is to use the Wget script provided automatically by the index node, in order to download more than one file from one data node. After the search and selection of the files, download and run the Wget script.

Wget script has a number of very useful features:

- recognize if files have already been downloaded and skip them;
- interrupted download before having finished can simply restarted by running the script in the same directory
- recognize if a new version of the downloaded data is available in ESGF

Wget script can also be created using a special URL, with help of the ESGF search RESTful API, so the search can be automatized e.g.

http://esgf-node.llnl.gov/search/cmip6/?source_id=EC-Earth3&experiment_id=dcppA-hindcast&sub_experiment=s1960&variable_id=tas&table_id=Amon

More information on the Wget script is available in the ESGF User Support Page.

There are additional very advanced ways to download data from ESGF like the semi-manual procedure using sprocket or the ESGF PyClient.

Finally, there are suites like ESMValTool where someone can download data from ESGF and in addition can use it as a diagnostics and performance metrics tool for the evaluation of Earth System Models (ESMs) that allows for routine comparison of single or multiple models. More information about ESMValTool can be found on the ESMValTool documentation page and on tutorial page

Appendix II C3S Recommendations for the Decadal Predictions

The C3S recommendations for the decadal predictions, formulated in the context of the Copernicus D34c contract, can be considered as an enhancement of the CMORized NetCDF data standards by adding standards for the search, identification, and dissemination of forecast data for multi-model systems. In that way, the usability of the data is improved when multiple overlapping experiments are used together. This will make the implementation of multi-model ensembles more straightforward.

The C3S enhancement requirements include:

- The files shall be CMORized;
- The files shall be in compliance with the CF conventions;
- The files shall be ESGF compliant;
- File names and directory tree should be as close as possible to the CMIP6 standards but for any case fully identifying the start date. A modification in accordance to the first set of recommendations is made here that fully identifies the start date, adding the start month in the name of the file. It consists in changing the actual DCPP standard from something like "tos_Omon_EC-Earth3_dcppA-hindcast_s1966-r1i1p1f1_gn_196611-196710.nc" to "tos_Omon_EC-Earth3_dcppA-hindcast_s196611-r1i1p1f1_gn_196611-196710.nc".
- Two additional time coordinates, one for the verification time (time) and another one for the forecast time (leadtime), should be included to ensure that the two time axes are fully documented. A scalar auxiliary coordinate (reftime) is needed for the start date. An important addition to the first set of recommendations is the requirement for the reftime and leadtime variables to have identical units, which should simplify the manipulation of the data.
  - "time" remains the only temporal dimension for the variables: tas(time,lat,lon) for example
  - The cell methods should take the double time axis into account when setting the corresponding bounds for non-instantaneous data.
  - Both reftime and leadtime should appear among the coordinates (in the CF sense of coordinates, attributes of the variable, eg "tas coordinates = height reftime leadtime")
  - of the physical variable to make sure that this important information is not lost during the data processing.
  - The global attributes 'sub_experiment', 'sub_experiment_id' and 'startdate' should be included and made compatible with the value of the auxiliary coordinate documenting the start date.
- The ensemble member is identified with the variable realization, an integer scalar coordinate with 'standard_name realization'. The variable takes the value of the initialization member coming from the ripf description, starting from 1.
- The realization variable should contain a comment attribute stating "For more information on the ripf, refer to the variant_label,

initialization_description,physics_description and forcing_description in the global

- attributes", where variant_label, initialization_index, initialization_description, physics_index,
- physics_description, forcing_index and forcing_description are global attributes that describe in more detail the CMIP6 ripf ensemble identifier.

In summary, the C3S enhancements require the following modifications:

- A "realization" variable is added, to represent the ensemble member
- The "sub_experiment_id" global attribute is adjusted to include the start year and month of the simulation
- A "reftime" variable is added, representing the start time of the simulation
- A "leadtime" coordinate variable is added: this is calculated from the "reftime" plus the valid times from the existing time variable

The C3S enhancements have been already implemented in the Decadal prediction data (DCPP) going to be published in CDS. The implementation for the decadal prediction is described in detail in the deliverable document C3S_34g_M3.1_202201_WPS_for_Decadal projections_v1 (can be obtained from ECMWF).

An example of a CMORized NetCDF following the C3S recommendation for decadal prediction is below:

```
netcdf tas_Amon_EC-Earth3_dcppA-hindcast_s198611-r1i1p1f1_gr_198611-198710 {
dimensions:
        time = UNLIMITED ; // (12 currently)
        lat = 256 ;
        lon = 512 ;
        bnds = 2 ;
variables:
        int reftime ;
                reftime:long_name = "Start date of the forecast" ;
                reftime:standard_name = "forecast_reference_time" ;
                reftime:units = "days since 1850-01-01" ;
                reftime:calendar = "gregorian" ;
        double leadtime(time) ;
                leadtime:long_name = "Time elapsed since the start of the forecast" ;
                leadtime:standard_name = "forecast_period" ;
                leadtime:units = "days" ;
        int realization ;
                realization:long_name = "realization" ;
                realization:comment = "For more information on the ripf, refer to the
variant_label, initialization_description, physics_description and forcing_description global
attributes" ;
        double time(time) ;
                time:bounds = "time_bnds" ;
                time:axis = "T" ;
                time:standard_name = "time" ;
                time:units = "days since 1850-01-01" ;
```

```
                time:calendar = "gregorian" ;
                time:long_name = "valid_time" ;
        double time_bnds(time, bnds) ;
                time_bnds:units = "days since 1850-01-01" ;
        double lat(lat) ;
                lat:bounds = "lat_bnds" ;
                lat:units = "degrees_north" ;
                lat:axis = "Y" ;
                lat:long_name = "latitude" ;
                lat:standard_name = "latitude" ;
        double lat_bnds(lat, bnds) ;
        double lon(lon) ;
                lon:bounds = "lon_bnds" ;
                lon:units = "degrees_east" ;
                lon:axis = "X" ;
                lon:long_name = "Longitude" ;
                lon:standard_name = "longitude" ;
        double lon_bnds(lon, bnds) ;
        double height ;
                height:units = "m" ;
                height:axis = "Z" ;
                height:positive = "up" ;
                height:long_name = "height" ;
                height:standard_name = "height" ;
        float tas(time, lat, lon) ;
                tas:standard_name = "air_temperature" ;
                tas:long_name = "Near-Surface Air Temperature" ;
                tas:comment = "near-surface (usually, 2 meter) air temperature" ;
                tas:units = "K" ;
                tas:cell_methods = "area: time: mean" ;
                tas:cell_measures = "area: areacella" ;
                tas:history = "2019-05-11T15:53:32Z altered by CMOR: Treated scalar dimension:
\'height\'. 2019-05-11T15:53:32Z altered by CMOR: Reordered dimensions, original order: lat lon
time." ;
                tas:missing_value = 1.e+20f ;
                tas:_FillValue = 1.e+20f ;
                tas:coordinates = "height reftime leadtime" ;

// global attributes:
                :Conventions = "CF-1.7 CMIP-6.2" ;
                :activity_id = "DCPP" ;
                :branch_method = "no parent" ;
                :branch_time = 0. ;
                :branch_time_in_child = 0. ;
                :branch_time_in_parent = 0. ;
                :contact = "cmip6-data@ec-earth.org" ;
                :creation_date = "2019-05-11T15:53:33Z" ;
                :data_specs_version = "01.00.27" ;
                :experiment = "hindcast initialized based on observations and using historical
forcing" ;
                :experiment_id = "dcppA-hindcast" ;
                :external_variables = "areacella" ;
                :forcing_index = 1 ;
                :frequency = "mon" ;
```

```
:further_info_url = "https://furtherinfo.es-doc.org/CMIP6.EC-Earth-Consortium.EC-Earth3.dcppA-hindcast.none.r1i1p1f1" ;
:grid = "ORCA1T255" ;
:grid_label = "gr" ;
:initialization_index = 1 ;
:institution = "AEMET, Spain; BSC, Spain; CNR-ISAC, Italy; DMI, Denmark; ENEA, Italy; FMI, Finland; Geomar, Germany; ICHEC, Ireland; ICTP, Italy; IDL, Portugal; IMAU, The Netherlands; IPMA, Portugal; KIT, Karlsruhe, Germany; KNMI, The Netherlands; Lund University, Sweden; Met Eireann, Ireland; NLeSC, The Netherlands; NTNU, Norway; Oxford University, UK; surfSARA, The Netherlands; SMHI, Sweden; Stockholm University, Sweden; Unite ASTR, Belgium; University College Dublin, Ireland; University of Bergen, Norway; University of Copenhagen, Denmark; University of Helsinki, Finland; University of Santiago de Compostela, Spain; Uppsala University, Sweden; Utrecht University, The Netherlands; Vrije Universiteit Amsterdam, the Netherlands; Wageningen University, The Netherlands. Mailing address: EC-Earth consortium,Rossby Center, Swedish Meteorological and Hydrological Institute/SMHI, SE-601 76 Norrkoping, Sweden" ;
:institution_id = "EC-Earth-Consortium" ;
:mip_era = "CMIP6" ;
:parent_activity_id = "no parent" ;
:parent_experiment_id = "no parent" ;
:parent_mip_era = "no parent" ;
:parent_source_id = "no parent" ;
:parent_sub_experiment_id = "no parent" ;
:parent_time_units = "no parent" ;
:parent_variant_label = "no parent" ;
:physics_index = 1 ;
:product = "model-output" ;
:realization_index = 1 ;
:realm = "atmos" ;
:source = "EC-Earth3 (2019): \n",
        "aerosol: none\n",
        "atmos: IFS cy36r4 (TL255, linearly reduced Gaussian grid equivalent to 512 x 256 longitude/latitude; 91 levels; top level 0.01 hPa)\n",
        "atmosChem: none\n",
        "land: HTESSEL (land surface scheme built in IFS)\n",
        "landIce: none\n",
        "ocean: NEMO3.6 (ORCA1 tripolar primarily 1 deg with meridional refinement down to 1/3 degree in the tropics; 362 x 292 longitude/latitude; 75 levels; top grid cell 0-1 m)\n",
        "ocnBgchem: none\n",
        "seaIce: LIM3" ;
:source_id = "EC-Earth3" ;
:source_type = "AOGCM" ;
:sub_experiment = "initialized near end of year 1986" ;
:table_id = "Amon" ;
:table_info = "Creation Date:(20 July 2018) MD5:b534c310c852aa1f0b00e68e90486e7a" ;
:title = "EC-Earth3 output prepared for CMIP6" ;
:tracking_id = "hdl:21.14100/b74313cc-6c6b-4b61-95f6-6d601b02bace" ;
:variable_id = "tas" ;
:variant_label = "r1i1p1f1" ;
:license = "CMIP6 model data produced by EC-Earth-Consortium is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (https://creativecommons.org/licenses). Consult https://pcmdi.llnl.gov/CMIP6/TermsOfUse for terms of use governing CMIP6 output, including citation requirements and proper acknowledgment. Further information about this data, including some limitations, can be found via the further_info_url (recorded as a global attribute in this file) . The data producers and data providers make no warranty, either express or implied, including, but not limited to, warranties of merchantability and fitness for a particular purpose. All liabilities arising from the supply
```

```
of the information (including any liability arising in negligence) are excluded to the fullest extent permitted by law." ;
                :cmor_version = "3.4.0" ;
                :variant_info = "Atmosphere initialization based on full-fields from ERA-Interim (s1979-s2018) or ERA-40 (s1960-s1978); ocean/sea-ice initialization based on full-fields from NEMO/LIM assimilation run nudged towards ORA-S4 (s1960-s2018)" ;
                :nominal_resolution = "100 km" ;
                :nco_openmp_thread_number = 1 ;
                :forcing_description = "Free text describing the forcings" ;
                :physics_description = "Free text describing the physics method" ;
                :initialization_description = "Free text describing the initialization method" ;
                :startdate = "s198611" ;
                :sub_experiment_id = "s198611" ;
data:

 reftime = "1986-10-31" ;

 leadtime = 16, 46, 76, 106, 136, 166, 196, 226, 256, 286, 316, 346 ;
}
```

Appendix III List of Variables

In the tables below, NetCDF DCPP and GRIB variables used in the project is provided.

Pressure Levels Fields

GRIB Name (C3S name)	GRIB short name	Units	GRIB ParamID	DCPP name	CF standard name	CMIP6 Short Name	CF canonical units
Geopotential	z	$m^2 s^{-2}$	129	geopotential	geopotential	zg, zg500	$m^2 s^{-2}$
Temperature	t	K	130	temp3d	air_temperature	ta	K
U-component of wind	u	$m s^{-1}$	131	ew wind	x_wind	ua	$m s^{-1}$
V-component of wind	v	$m s^{-1}$	132	ns wind	y_wind	va	$m s^{-1}$
Specific humidity	q	$kg kg^{-1}$	133	spec humidity	specific_humidity	hus	1

Single-level parameters

GRIB Name (C3S name)	GRIB short name	Units	GRIB ParamID	DCPP name	CF standard name	CMIP6 Short Name	CF canonical units
2 metre temperature	2t	K	167	sfc air T	air_temperature	tas	K
2 metre dewpoint temperature	2d	K	168	dew point temp	dew_point_temperature	tdps	K
10 metre U wind component	10u	$m s^{-1}$	165	ew wind	x_wind	uas	$m s^{-1}$
10 metre V wind component	10v	$m s^{-1}$	166	ns wind	x_wind	vas	$m s^{-1}$
-	-	-	-	day mean wind	wind_speed	sfcWind	$m s^{-1}$

-	-	-	-	specific humidity	specific_humidity	huss	1
	-	-	-	relative humidity (CMIP6 variable)	relative_humidity	hurs	%
	-	-	-	Daily minimum relative humidity (CMIP6 variable)	relative_humidity	hursmin	%
Mean sea level pressure	msl	Pa	151	mean sea level pressure	air_pressure_at_sea_level	psl	Pa
Surface pressure	sp	Pa	134	sfc pres	surface_air_pressure	ps	Pa
Total cloud cover	tcc	(0-1)	164	cld frac	cloud_area_fraction	clt	1
Soil temperature (level1)	stl1	K	139	-	soil_temperature		K
Skin temperature	skt	K	235	skin temp	surface_temperature	ts	K
Volumetric soil moisture	vsw	m^{-3} m^{-3}	260199	soil moisture	mass_content_of_water_in_soil	mrso	kg m^{-2}
				frozen soil moisture	soil_frozen_water_content	mrfso	kg m^{-2}
Sea surface temperature	sst	K	34	sst	sea_surface_temperature	tso (C3S short name)	K
	-	-	-	sst	sea_surface_temperature	tos	K
Sea ice area fraction Sea-ice cover	ci	(0-1)	31	ice fraction	sea_ice_area_fraction	sic	1
	-	-	-	ice thinkmess	sea_ice_thickness	sit	m
Snow depth	sd	m of water equivalent	141	-	lwe_thickness_of_surface_snow_amount	-	m

Snow density	rsn	kg m^{-3}	33	snow density	snow_density	-	kg m^{-3}
Maximum temperature at 2 metres in the last 24 hours	mx2t24	51	K	day t max	air_temperature	tasmax	K
Minimum temperature at 2 metres in the last 24 hours	mn2t24	52	K	day t min	air_temperature	tasmin	K
10 metre wind gust since previous post-processing	10fg	49	m s^{-1}	day max wind	wind_speed_of_gust wind_speed	sfcWindmax	m s^{-1}
Total precipitation Large-scale precipitation Convective precipitation	tp (total) lsp (large scale) cp (convective)	228 (total) 142 (large scale) 143 (convective)	m	-	lwe_thickness_of_precipitation_amount lwe_thickness_of_stratiform_precipitation_amount lwe_thickness_of_convective_precipitation_amount	-	m m m
	-	-	-	net pcp	precipitation_flux	pr	kg m-2 s-1
	-	-	-	day max houtly pcp	precipitation_flux	pr	kg m-2 s-1
Snowfall	sf	144	m of water equivalent	-	lwe_thickness_of_snowfall_amount	-	m
	-	-	-	pcp as snow	snowfall_flux	prsn	kg m-2 s-1
Surface sensible heat flux	sshf	146	J m^{-2}	sensible up	surface_upward_sensible_heat_flux	hfss	W m^{-2}
Surface latent heat flux	slhf	147	J m-2	latent up	surface_upward_latent_heat_flux	hfls	W m^{-2}

Surface solar radiation downwards	ssrd	169	J m-2	solar down	surface_downwelling_shortwave_flux_in_air	rsds	W m^{-2}
Surface thermal radiation downwards	strd	175	J m-2	lw down	surface_downwelling_longwave_flux_in_air	rlds	W m^{-2}
Surface net solar radiation	ssr	176	J m-2	net solar	surface_net_downward_shortwave_flux	rss	W m^{-2}
Surface net thermal radiation	str	177	J m-2	net lw	surface_net_downward_longwave_flux	rls	W m^{-2}
Top net solar radiation	tsr	178	J m-2	-	toa_net_downward_shortwave_flux	-	W m^{-2}
TOA incident solar radiation Top incoming solar radiation	tisr	212	J m-2	toa solar incident	toa_incoming_shortwave_flux	rsdt	W m^{-2}
-	-	-	-	toa solar out	toa_outgoing_shortwave_flux	rsut	W m^{-2}
-	-	-	-	toa lw out	toa_outgoing_longwave_flux	rlut	W m^{-2}
Top net thermal radiation	ttr	179	J m-2	toa lw out	toa_net_downward_longwave_flux	rlut	W m^{-2}
	-	-	-		surface_upwelling_shortwave_flux_in_air	rsus	W m^{-2}

					surface_upwelling_longwave_flux_in_air	rlus	W m^{-2}
Eastward turbulent surface stress	ewss	180	N m^{-2} s	ew stress down	surface_downward_eastward_stress	tauu	Pa
Northward turbulent surface stress	nwss	181	N m^{-2} s	ns stress down	surface_downward_northward_stress	tauv	Pa
Evaporation	e	182	m of water equivalent	-	lwe_thickness_of_water_evaporation_amount	-	m
	-	-	-	net evap	water_evapotranspiration_flux	evspsbl	kg m-2 s
Runoff (or surface & sub-surface runoff)	ro (total) sro (Surface runoff) ssro (Sub-surface runoff)	205 (total) 8 (Surface runoff) 9 (Sub-surface runoff)	m	runoff	runoff_amount surface_runoff_amount subsurface_runoff_amount	mrro	kg m^{-2}
Runoff	ro	205	m	runoff	runoff_flux	mrro	kg m-2 s-1
Depth average potential temperature of upper 300m	thetao	151129	degC	depth avg pot temp (300, 700, 2000m)	sea_water_potential_temperature	thetao	K
Depth average salinity of upper 300m	so	262500	10^{-3}	mean salinity in upper 300m	sea_water_practical_salinity	so	psu
				ew speed	sea_water_x_velocity	uo	m s^{-1}

				ns speed	sea_water_x_velocity	vo	m s^{-1}
Depth of 14°C isotherm	t14d	262102	m	depth 14c	depth_of_isosurface_of_sea_water_potential_temperature	t14d	m
Depth of 17°C isotherm	t17d	262103	m	depth 17c	depth_of_isosurface_of_sea_water_potential_temperature	t17d	m
Depth of 20°C isotherm	t20d	262104	m	depth 20c	depth_of_isosurface_of_sea_water_potential_temperature	t20d	m
Depth of 26°C isotherm	t26d	262105	m	depth 26c	depth_of_isosurface_of_sea_water_potential_temperature	t26d	m
Depth of 28°C isotherm	t28d	262106	m	depth 28c	depth_of_isosurface_of_sea_water_potential_temperature	t28d	m
Mixed layer depth 0.01	mlotst010	151225	m	thickness mix layer	ocean_mixed_layer_thickness_defined_by_sigma_theta	mlotst	m
Mixed layer depth 0.03	mlotst0300	151226	m		ocean_mixed_layer_thickness_defined_by_sigma_theta		m
Sea ice thickness (sea ice volume)	sithick	174098	m		sea_ice_thickness	sivol	m
Sea surface height above geoid	zos	151145	m	sea sfc height	sea_surface_height_above_geoid	zos	m

Sea surface salinity	sos	151219		-	sea_water_practical_salinity	-	psu
	-	-	-	sss	sea_surface_salinity	sos	10-3
	-	-	-	thickness mix layer	ocean_mixed_layer_thickness_defined_by_sigma_t	mlotst	m
	-	-	-		ocean_meridional_overturning_mass_streamfunction	msftmyz	kg s^{-1}
Orography (geopotential height)	orog	156	gpm		surface_altitude		1
Land area fraction	sftlf	Categorical (0-1)	172		land_area_fraction		m