Copernicus Climate Data Store
Data and Metadata QC
(C3S_512)

RDA Spain webinar 2020, 9 July

Joan Sala Calero
joan.sala@bsc.es
Pierre-Antoine Bretonnière
pierre-antoine.bretonnière@bsc.es
Introduction: COPERNICUS CLIMATE CHANGE SERVICE (C3S)

- Copernicus EO programme (EC)
- C3S will provide past/present/future climate data
- To a wide range of users
- The goal is to achieve a more sustainable future
At the heart of the C3S infrastructure is the **Climate Data Store (CDS)**. It provides a single point of access to a wide range of climate datasets, namely satellite and in-situ observations, reanalyses, seasonal forecasts and climate projections.

The users access the datasets through the CDS portal with the option of:

- Downloading,
- Visualizing,
- Subsetting,
- Plotting,
- Etc.

Data and Metadata challenges: C3S_512 contract

Such a complex infrastructure requires an Evaluation and Quality Control (EQC) function providing an overarching quality assurance service for the whole CDS data and metadata:

- Apply **FAIR** principles to **metadata**
  - Use of ISO 19115 metadata record standard
  - **Interoperable**
  - OGC services such as CSW (Catalogue Service)
  - **Findable/Accessible**
  - Unique identifiers for the data products / Use of DOIs
  - **Findable/Accessible**
  - Global and European data standards (WMO OAI-PMH / EU INSPIRE)
  - **Interoperable**

- Apply **quality controls** to all the **data** provided to the users
  - Use of meteorological file formats (GRIBv2, NetCDF) - **Interoperable/Reusable**
  - Based on standard climate vocabularies (CF conventions, CMOR)
  - **Interoperable**
  - Plausible/Valid ranges for the data we provide
  - Spatial/Temporal completeness and consistency

---

ERAS monthly averaged data on single levels from 1979 to present

- ERAS is the fifth generation ECMWF reanalysis for the global climate and weather for the past 4 to 7 decades. Currently data is available from 1979. When complete, ERAS will contain a detailed record from 1950 onwards. ERAS replaces the ERA-40 reanalysis.
- Reanalysis combines observations into globally complete fields using the laws of physics with the method of data assimilation. In the case of ERA-40, ERA-40 produces a dynamical estimate for a large number of atmospheric, oceanic, and land-surface quantities. An uncertainty estimate is generated by an ensemble of 15 member analyses at three-hourly intervals. Ensemble mean and spread have been pre-computed for convenience. Such uncertainty estimates are closely related to the information content of the available observing system which has evolved dramatically over time. They also indicate how dependent certain areas are.
- The name resolution of the ERA-40 atmosphere and land reanalysis is 21 km (orographic grid (10 km) and 3 km (1.5 km) for the ensemble members. Oceanic products are produced at 0.5 degrees and 1 degree for the ensemble. The ensembles have a component of 170 levels in the vertical from the surface up to 1 (1.5) Kelvin.
- This covers the troposphere, stratosphere and mesosphere. There are both analyses fields and short forecast fields. Two analysis versions exist at 40-km. A detailed description can be found in the online ERAS documentation. The full data set resides in the ERAS tape archive.
- The data presented here is a post-processing output of the full ERA-40 data set. It is online in spin-up, which should ensure fast and easy access. It should satisfy the requirements for most common applications.
- Data has been regrided to a regular latitude grid of 0.25 degrees for the reanalysis and from 0.5 degrees for the uncertainty estimate 0.25 and 1 degree respectively for ocean winds. There are two main sub sets: data on pressure levels and data on single levels. The data on pressure levels contain 10 atmospheric quantities on 35 pressure levels from 1,000 hPa (equivalent to 1 NPa on the top of the stratosphere). Single-level data are available for a number of atmospheric, oceanic and land surface quantities.
Work performed during the RDA 4.0 grant

- Standardization of the vocabularies used in the metadata using the CF (Climate and Forecast) conventions.
- Use of the Common Information Model (CIM) to describe:
  - Data, models and software used to produce it.
  - Geographic grids and projections.
  - Experiments or simulations.
- Development of an INSPIRE KPI using the REST API of the inspire validator focusing on the CSW metadata catalogue: [https://inspire.ec.europa.eu/validator/](https://inspire.ec.europa.eu/validator/)
Dissemination activities performed during the RDA 4.0 grant

- Presentation during the EOSC symposium 2019 (Budapest) in the use cases session.

- Publication on the EOSC portal use cases in the coming months: https://www.eosc-portal.eu/eosc-in-practice/use-cases

- RDA adoption story will be online soon: https://www.rd-alliance.org/recommendations-outputs/adoption-stories
THANK YOU

The research leading to these results is part of the Copernicus Climate Change Service (C3S)

(Service Contract number: ECMWF/COPERNICUS/2018/C3S_512_BSC), a programme being implemented by the European Centre for Medium-Range Weather Forecasts (ECMWF) on behalf of the European Commission.

joan.sala@bsc.es
pierre-antoine.bretonniere@bsc.es