

# ERA-CLIM2 Review M27

## WP5: Service Developments

Manuel Fuentes

Products Team – Production Section - ECMWF



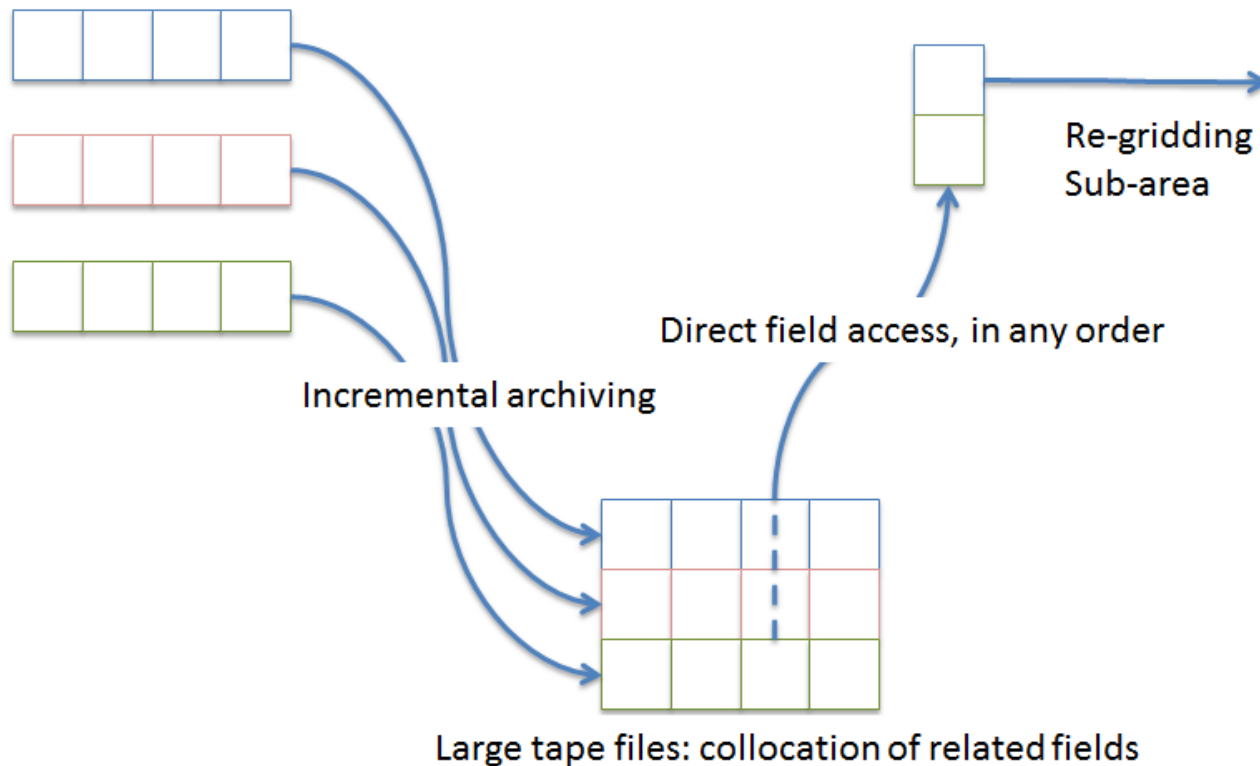
## WP5 Deliverables

**D5.1** : Technical developments in MARS to support **archiving and retrieval of data in NetCDF** format. This is needed to allow use of MARS for the **ocean component** of the coupled reanalysis output. [**month 30-36**]

**D5.2** : Implementation of **public data services** for gridded output from coupled climate reanalyses (CERA), including the ocean component. [**month 48**]

**D5.3** : **Report** on data services usage and **requirements** for climate reanalysis [**month 48**]

# Service offered by MARS



- MARS archives atmospheric field of CERA-20C (handles 2D fields in GRIB)
- MARS scans archived files, extracts metadata from GRIB headers and keeps an index that tracks where each GRIB field is
- GRIBs are reorganised into larger files, to minimise the total number of files and collocate related fields to speed up retrievals
- On retrieval, MARS find the required fields, reads them from tape, and re-assembles them according to the user's request

# Production of CERA-20C



## Navigation

[Public Datasets](#)

[Job list](#)

[< Return to selection](#)

## Additional filtering

### Current request

*Stream:* Atmospheric model

*Parameter:* 10 metre U wind component, 10 metre V wind component, 100 metre U wind component, 100 metre V wind component, 2 metre dewpoint temperature, 2 metre temperature

*Year:* 1900

*Number:* 0 to 9

*Month:* February

*Version:* 2366

*Type of level:* Surface

*Time:* 00:00:00, 03:00:00, 06:00:00, 09:00:00, 12:00:00, 15:00:00, 18:00:00, 21:00:00

*Date:* 19000203

*Type:* Analysis

*Class:* ERA-CLIM2 coupled reanalysis of the 20th-century (CERA-20C)

### The request will be done using the following attributes:

**Area:** Default (as archived) ([change](#))

**Grid:** 1x1 ([change](#))

## D5.1 Support for NetCDF in MARS

Original NEMO output files contain many variables (2D, 3D), feedback files, restart files, ocean observations, all annotated with NEMO's own convention

### Solution considered:

- NetCDF files are split into individual NetCDF files, 2D or 3D
  - Resulting NetCDF files must follow an agreed convention based on CF
  - Resulting NetCDF files are annotated with MARS specific information. These attributes are used by MARS to index the NetCDF files, and treat them as simple binary records
- On retrieval, those records will be assembled in a single NetCDF file to be delivered to the user

### Challenges:

- Define what variables to archive, focusing on user service
  - not all output above is interesting to users
- Find CF standard names
  - About 60% variables have a CF standard name
- Find sound metadata to enable assembling records on retrieval

# Prototype of NetCDF Ocean output in MARS



## Navigation

[Public Datasets](#)

[Job list](#)

[< Return to selection](#)

## Additional filtering

### Current request

*Stream:* Atmospheric model

*Parameter:* Ocean salinity

*Year:* 1900

*Number:* 1

*Month:* February

*Version:* 2366

*Type of level:* Depth

*Time:* 00:00:00, 03:00:00, 06:00:00, 09:00:00, 12:00:00, 15:00:00, 18:00:00, 21:00:00

*Date:* 19000203

*Type:* Analysis

*Class:* ERA-CLIM2 coupled reanalysis of the 20th-century (CERA-20C)

### The request will be done using the following attributes:

**Area:** Default (as archived) ([change](#))

**Grid:** 1x1 ([change](#))

# Support for NetCDF at ECMWF

- Other projects require NetCDF support
  - All data being served from ECMWF Data Portals
  - Sub-seasonal to Seasonal project (S2S) requires archiving in MARS of Ocean output from 11 production Centres (near real-time + reforecast)
  - Other forecasting systems producing ocean output (HRES, ENS, ....)
  - Seasonal forecast being delivered in NetCDF to the C3S (real-time + reforecast), and will need to be served by the Climate Data Store (CDS)
  - Climate Predictions (CMIP6) will be part of the C3S, will be served by the CDS
- Define common metadata that will allow tools to seamlessly handle all the above
  - Define a MARS/NetCDF convention, on top of CF, CMIP5/6, SPECS
  
- ERA-CLIM2 D5.1 will enable support for NetCDF at ECMWF
  - Transfer the development outcomes into preoperational phase of C3S

## D5.2 Public Data Services

### Consolidation of CERA-20C into user version

- After production is complete and quality of data has been checked, the various CERA-20C streams will be consolidated into a single version:
  - Retrieve all relevant data (GRIB/NetCDF Fields, ODB feedback)
  - Re-badge headers with the user version (0001)
  - Archive back into MARS
- This process is expected to take several months

### Public Data Server will be extended

- Public interface to data stored in MARS
- Expected to take less than 1 month after consolidation phase





# CERA Public Data Server: Observation Feedback

Navigation

- Datasets
- Job list
- Batch access

[Return to selection](#)

## retrieve

**Request abstract:**

Atmospheric model, DRIBU Ocean Bottle And Low Resolution Conductivity Temperature Depth CTD And XCTD ..., 1607, ERA-CLIM pilot reanalysis of the 20th-century using surface observations only, select reporttype, varno, timeseries\_index, date, time, lat, lon, obsvalue, fg\_depar, an\_depar where (source='ISPDv2.2') and (varno=110); all, 1994-04-01...2010-12-31, ODB feedback, 17

The status of the **job** is: **complete**

**Acknowledgement**

Support for the **International Surface Pressure Databank** is provided by the *U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE)* program, and *Office of Biological and Environmental Research (BER)*, and by the *National Oceanic and Atmospheric Administration (NOAA) Climate Program Office*. **The Twentieth Century Reanalysis Project** is supported by the *Earth System Research Laboratory Physical Sciences Division of NOAA* and the *Climate Diagnostics Center (CDC) of the University of Colorado's Cooperative Institute for Research in Environmental Sciences (CIRES)*.

**Download (9627.8MB)**

**Preview of the data**

reporttype	varno	timeseries_index	date	time	lat	lon	obsvalue	fg_depar	an_depar
16005	110	66	19940331	220000	52.500000	-132.699997	102270.000000	137.000000	78.300003
16005	110	1767	19940331	220000	39.290001	164.649994	101340.000000	-426.000000	-248.000000
16005	110	244	19940331	220000	-39.040001	163.289993	102440.000000	129.000000	53.299999
16005	110	2309	19940331	220000	20.639999	135.789993	101580.000000	168.000000	143.000000
16005	110	1255	19940331	220000	23.540001	141.110001	101980.000000	115.000000	88.000000
16005	110	4542	19940331	220000	9.480000	159.669998	101200.000000	57.799999	63.700001
16005	110	3745	19940331	220000	-8.140000	114.400002	101120.000000	-4.070000	-12.700000
16005	110	5060	19940331	220000	-12.380000	130.669998	101220.000000	91.900002	28.799999
16005	110	2512	19940331	220000	-11.780000	118.989998	101240.000000	55.599998	28.200001
16005	110	1751	19940331	220000	-13.360000	139.389999	101190.000000	78.500000	39.299999

# CERA Public Data Server: ECMWF WebAPI

- Simple API to services using HTTPS
  - Install a simple library + token
  - Download data via batch scripts
- Access to new services, like plots on demand

```
usr/bin/env python
```

```
from ecmwfapi import ECMWFDataServer
```

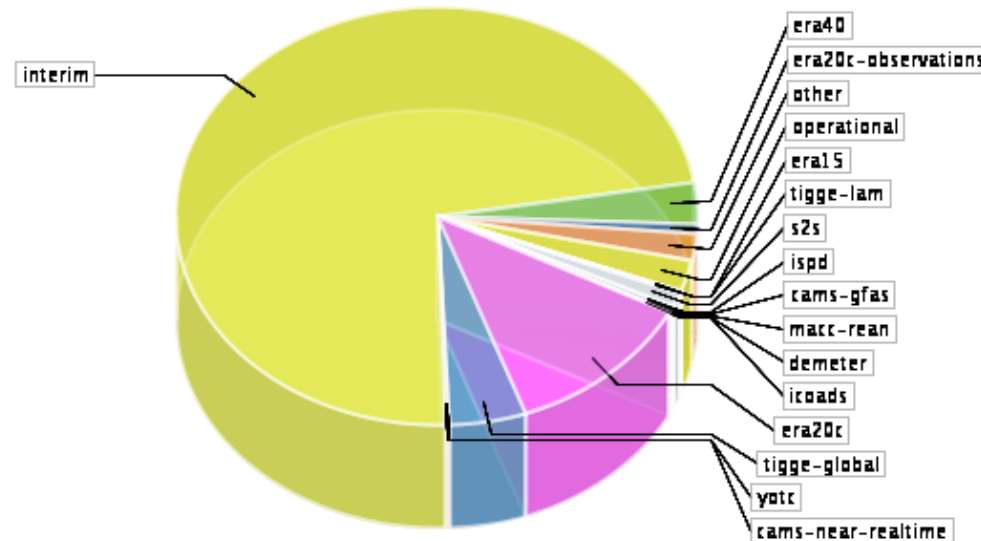
```
server = ECMWFDataServer()
```

```
server.retrieve({  
    'dataset' : "era20c",  
    'levtype' : "sfc",  
    'date'    : "20100101/to/20101231",  
    'time'    : "00",  
    'param'   : "2t",  
    'grid'    : "1/1",  
    'format'  : "netcdf",  
    'target'  : "data.nc"  
})
```

## D5.3 Report on data services usage

- Statistics on Public Data Servers under development
  - Enter all requests in a database, for further analysis
- Eg, total number of users per month in 2015 of the ERA-20C dataset

Total nr of requests



tigge-global (2,959,458 - 5%)	yotc (74,438 - 0%)	cams-near-realtime (98,045 - 0%)	interim (44,914,412 - 73%)	era40 (2,019,581 - 3%)
era20c-observations (465,052 - 1%)	other (1,264,953 - 2%)	operational (1,514,031 - 2%)	era15 (7,825 - 0%)	tigge-lam (5,908 - 0%)
s2s (744,017 - 1%)	ispd (44,002 - 0%)	cams-gfas (31,592 - 0%)	macc-rean (169,053 - 0%)	demeter (10,722 - 0%)
era20c (7,276,820 - 12%)	icoads (74,985 - 0%)			

# WP5: Service Development - Summary

- WP5 Deliverables on track:
  - D5.1 MARS support for NetCDF being finalised
    - Expected Q3 2016 (M33)
  - D5.2 CERA Data Servers: extend current infrastructure and services
    - CERA-20C Expected Q4 2016 (M36)
  - D5.3 Data services usage
    - Expected 2017 (M42-48)



Thank you for your attention!