Argo Data Management System

Argo Data Team, Thierry Carval

First DBCP Mediterranean Training Workshop on Ocean Observations and Data Applications

Coriolis

Argo

What is Argo

- Argo is an international program that collects information from **inside** the ocean using a fleet of profiling floats that drift with the ocean currents.
  https://argo.ucsd.edu

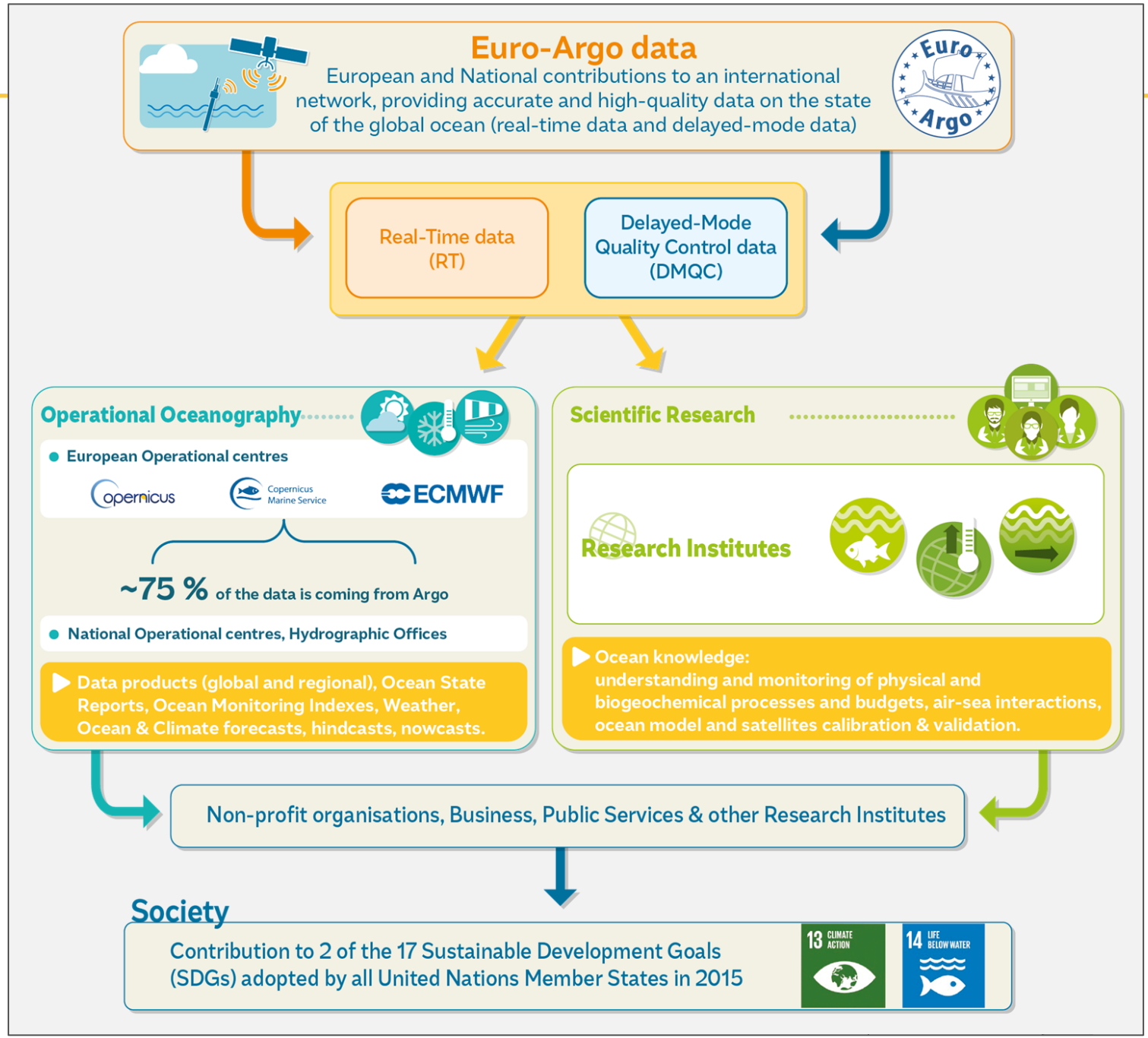
- The standard Argo float mission is a 10-day cycle, with most of the float's time spent drifting along with deep ocean currents, followed by taking a series of measurements as it moves back up (profiles) to the ocean surface.
  https://argo.ucsd.edu/how-do-floats-work

- Argo data policy is to deliver unrestricted data quality controlled in real-time and delayed mode from the 17 000 floats (3700 active)
  https://argo.ucsd.edu/data/data-from-gdacs

Argo data use

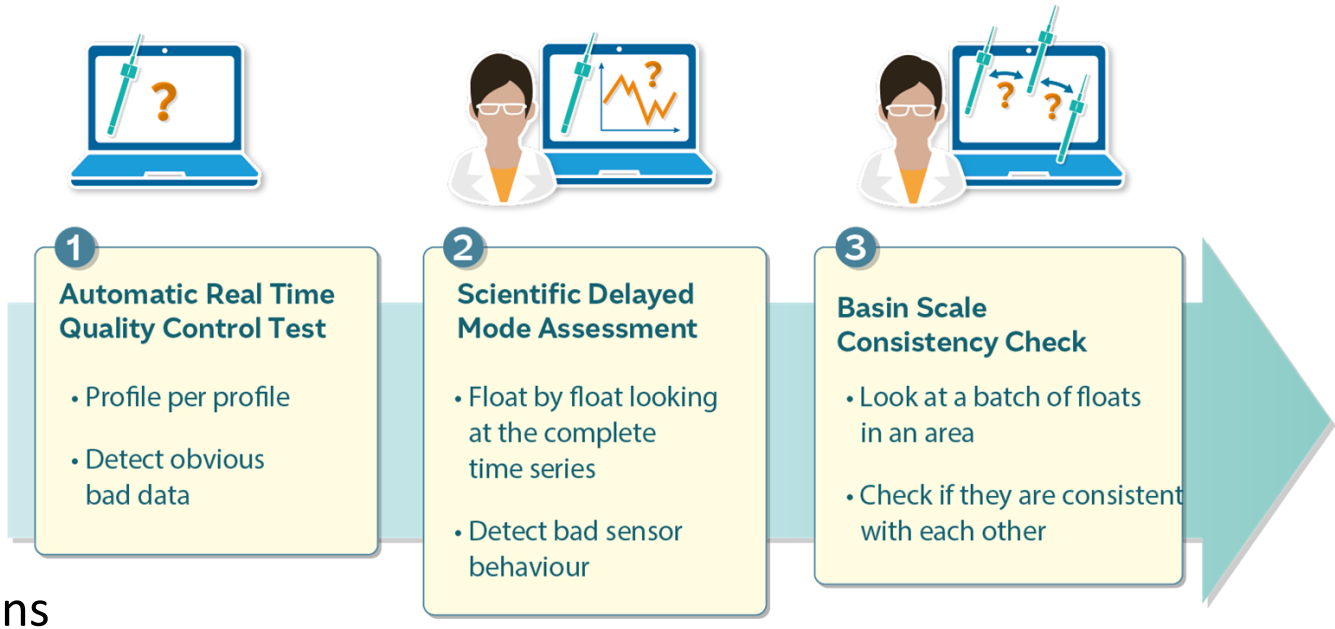
- Scientific research
- Operational services
  - Main ocean in situ data provider, complementary to satellite

Free & open data policy



2 data flows

- Real Time
  - ✓ 12 hours max
  - ✓ Automatic Quality Control tests
  - ✓ Operational applications

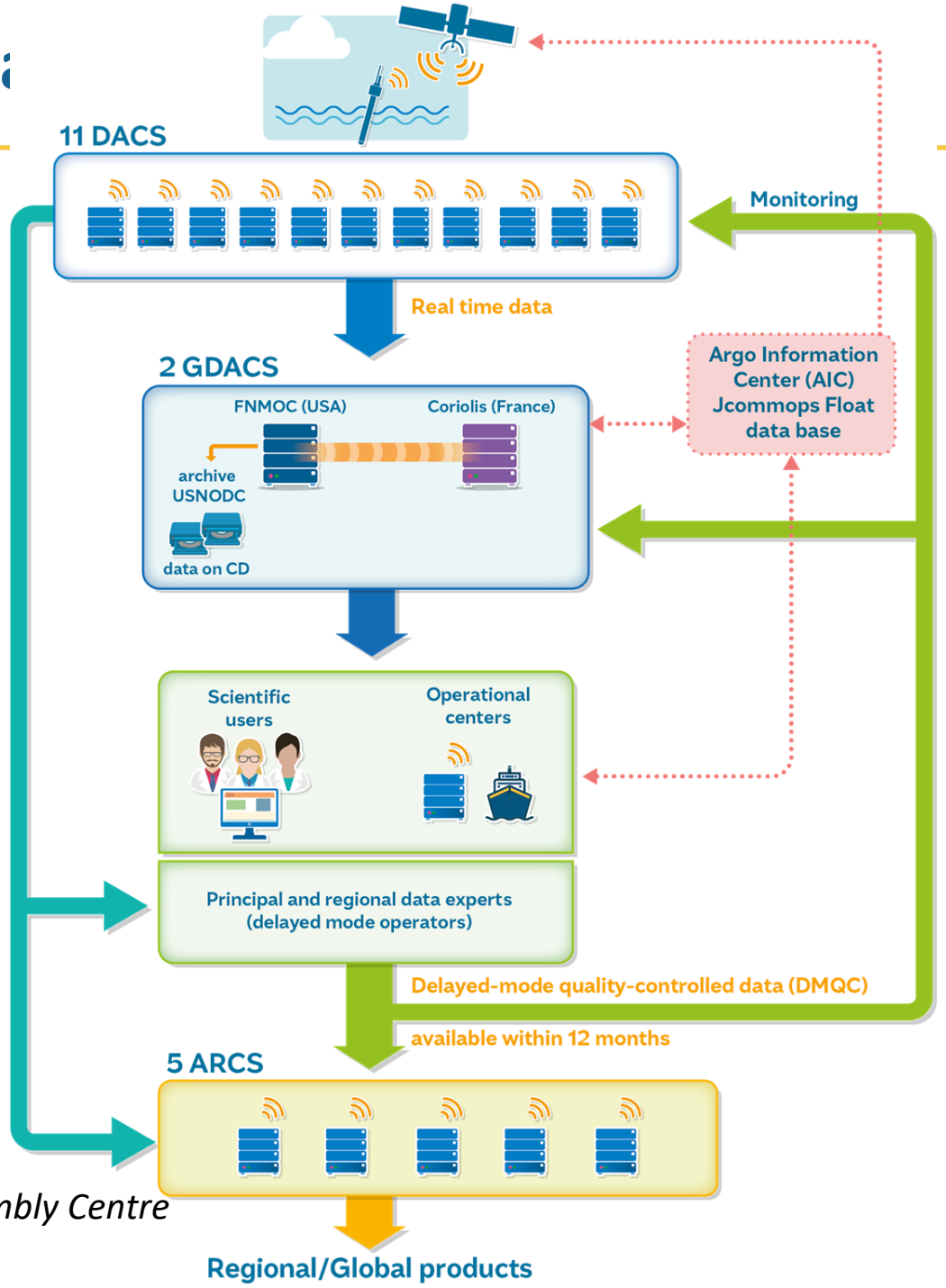


- Delayed Mode
  - ✓ 12 months
  - ✓ Detailed time serie analysis and corrections
  - ✓ Ocean & climate science applications

+ additional analysis at bassin scales (Argo Regional Centres)

Argo data system: managed at interna

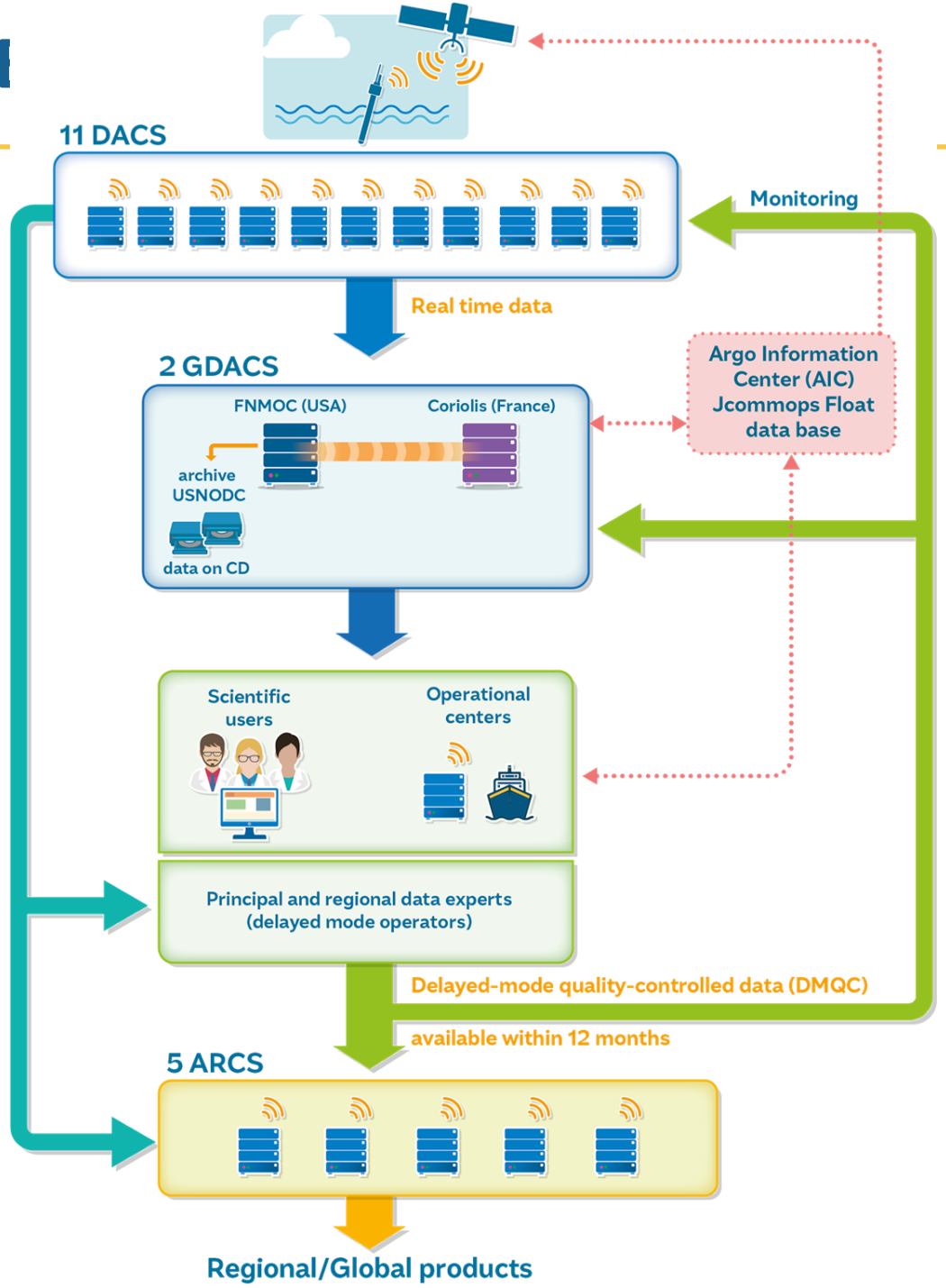
- Floats send their measurements to DACs*, where raw data are processed and sent to the 2 GDACs*
  - 2 GDAC : Coriolis/France & FNMOC/USA
  - 11 DACs worldwilde
  - 5 Argo Regional Centers
  - Argo Information Centre (AIC) at OceanOPS
    - ✓ Registration of floats (IOC resolution XX-6)
    - ✓ Information on data (“metadata”)



**(G)DAC* = *(Global) Data Assembly Centre*
**ARC* = *Argo Regional Centre*

Argo data system: Evolve to manage I

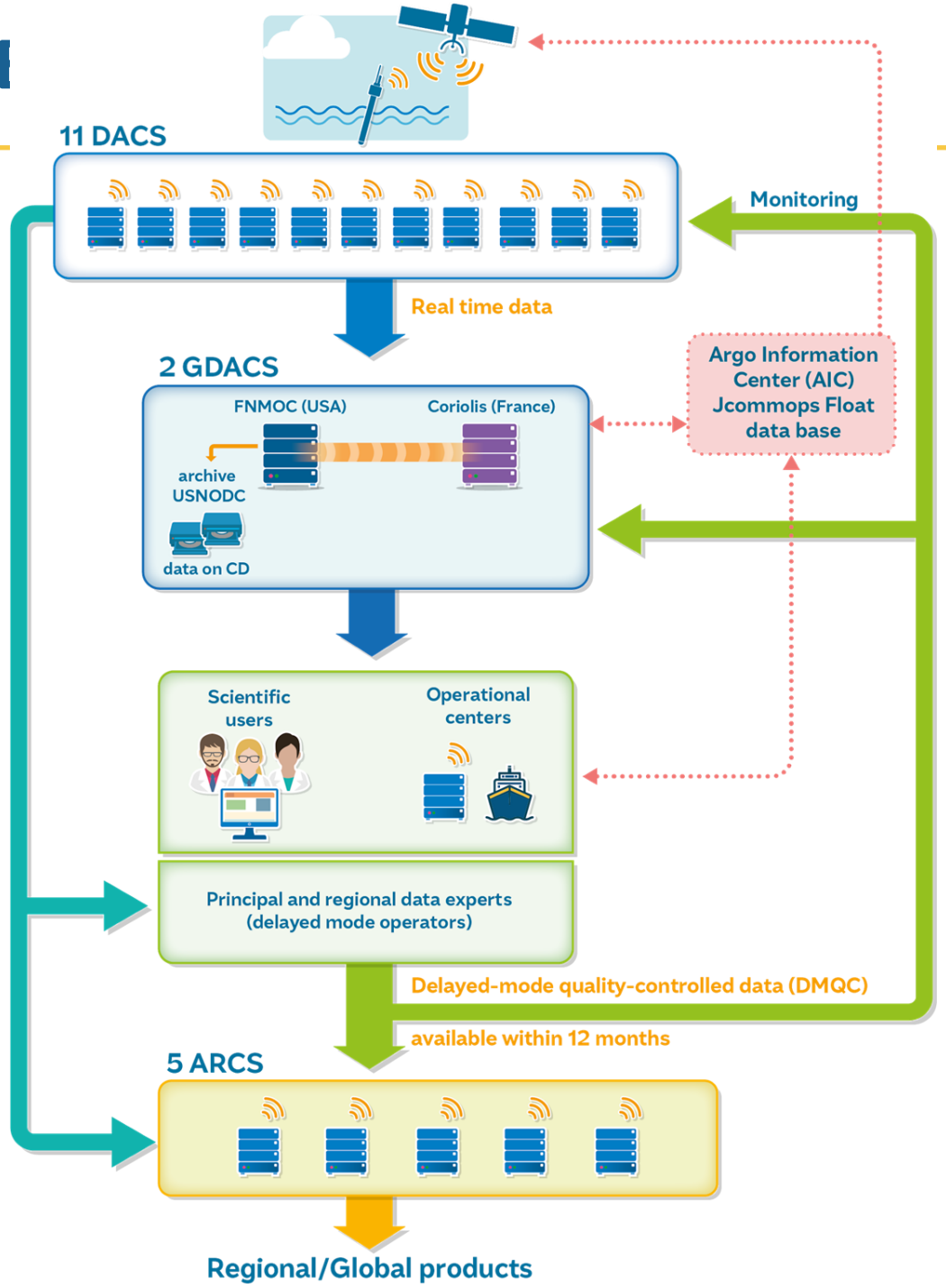
- Argo Data system developed for Temperature/Salinity/Pressure is extended to manage BGC-Argo parameters (Oxygen, Chlorophyll, Backscatter, Nitrate, pH, Irradiance)
  - Extension of Argo vocabularies
  - Enhancement of Argo data format
  - Development of Real-Time Quality Control for the 6 variables and implementation at DACs
  - Development of Delayed mode process Control for the 6 variables
  - Sharing of tools on collaborative platforms
- BGC Argo data system requires additional man power and expertise to reach the Ocean health and climate challenges
  - Delayed mode data processing needs to be funded and organized at international level



Argo data system: Evolve to manage I

- Real time data stream for BGC is more challenging than more P/T/S
  - Very few P/T/S need adjustment in real time at the early stage of their life time
  - More or less all of the BGC data needs to be adjusted in real time
  - Automated Real-Time adjustment for all BGC variables is still under development as we need enough BGC data at sea to build robust procedures to be implemented at DACs

⚠ Be very cautious with raw BGC data and focus on adjusted or DM data



Argo Data Management Team (ADMT)

- International team
  - DACs representatives & scientists
  - BGC Argo Data Management Task Team
- Meet annually during one full week:
  - to discuss Argo data related issues
  - to agree on data formats, data corrections methods and quality control procedures

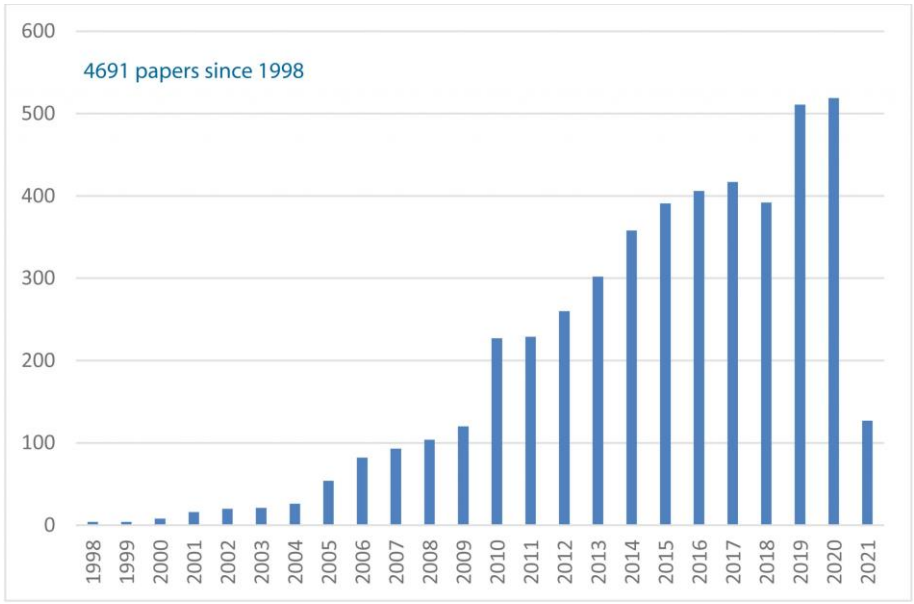
20th ADMT meeting participants
Villefranche-sur-Mer, France, October 2019

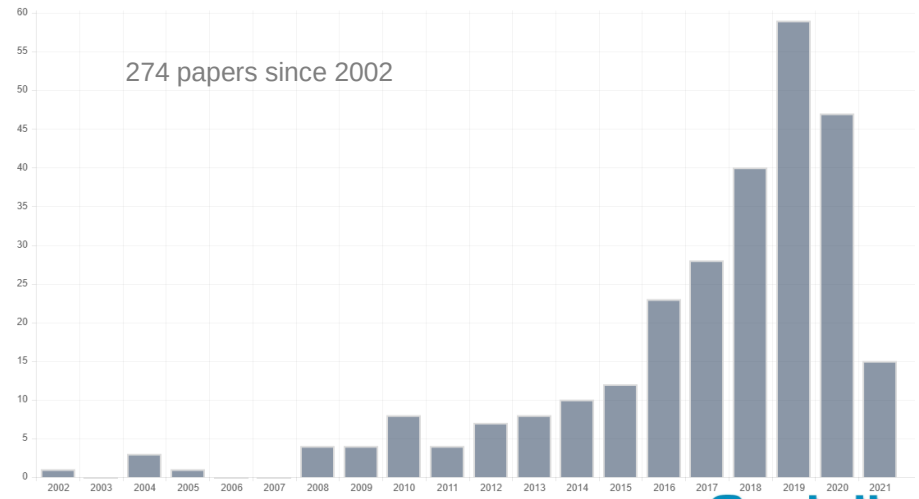
Argo data use: scientific research

Aknowledgement:

" These data were collected and made freely available by the International Argo Program and the national programs that contribute to it. (https://argo.ucsd.edu, https://www.ocean-ops.org). The Argo Program is part of the Global Ocean Observing System. "

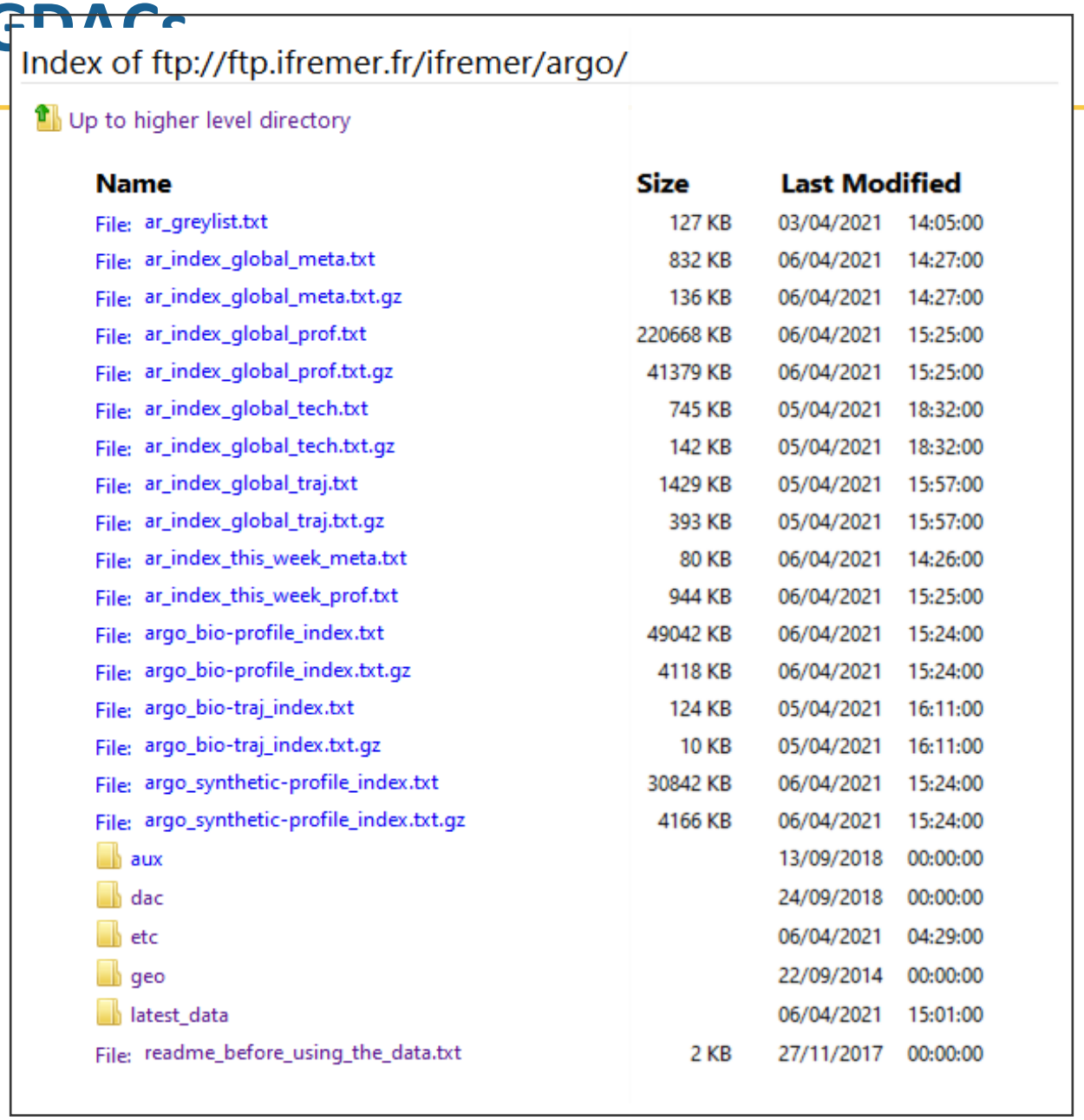
Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). https://doi.org/10.17882/42182





Argo data organisation on the GDACs

- GDACs are organized into three main folders:
  - a “**dac**” folder which sorts the data by Data Assembly Centre (DAC)
  - a “**geo**” folder which sorts the data by ocean basin
  - a “**latest_data**” folder which includes the most recent data
  - \+ “**aux**” folder which contains data from experimental sensors.
- Index files containing a list of metadata on each type of Argo data file (meta, prof, tech and traj) contained in the “dac” and “geo” folders.
- Greylist file which contains a list of floats that likely have sensor problems.



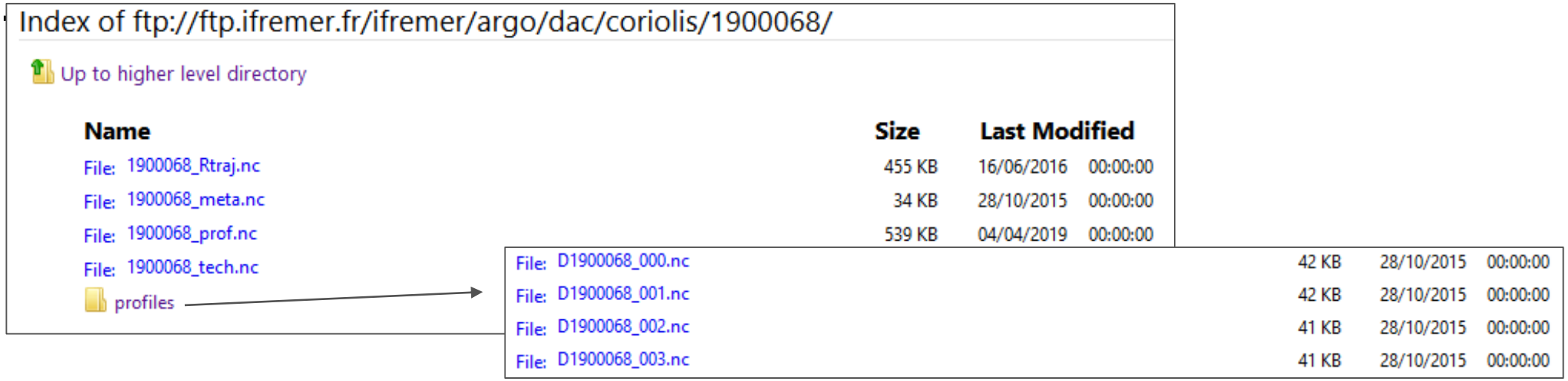
Index of ftp://ftp.ifremer.fr/ifremer/argo/

Up to higher level directory

Name	Size	Last Modified
File: ar_greylist.txt	127 KB	03/04/2021 14:05:00
File: ar_index_global_meta.txt	832 KB	06/04/2021 14:27:00
File: ar_index_global_meta.txt.gz	136 KB	06/04/2021 14:27:00
File: ar_index_global_prof.txt	220668 KB	06/04/2021 15:25:00
File: ar_index_global_prof.txt.gz	41379 KB	06/04/2021 15:25:00
File: ar_index_global_tech.txt	745 KB	05/04/2021 18:32:00
File: ar_index_global_tech.txt.gz	142 KB	05/04/2021 18:32:00
File: ar_index_global_traj.txt	1429 KB	05/04/2021 15:57:00
File: ar_index_global_traj.txt.gz	393 KB	05/04/2021 15:57:00
File: ar_index_this_week_meta.txt	80 KB	06/04/2021 14:26:00
File: ar_index_this_week_prof.txt	944 KB	06/04/2021 15:25:00
File: argo_bio-profile_index.txt	49042 KB	06/04/2021 15:24:00
File: argo_bio-profile_index.txt.gz	4118 KB	06/04/2021 15:24:00
File: argo_bio-traj_index.txt	124 KB	05/04/2021 16:11:00
File: argo_bio-traj_index.txt.gz	10 KB	05/04/2021 16:11:00
File: argo_synthetic-profile_index.txt	30842 KB	06/04/2021 15:24:00
File: argo_synthetic-profile_index.txt.gz	4166 KB	06/04/2021 15:24:00
aux		13/09/2018 00:00:00
dac		24/09/2018 00:00:00
etc		06/04/2021 04:29:00
geo		22/09/2014 00:00:00
latest_data		06/04/2021 15:01:00
File: readme_before_using_the_data.txt	2 KB	27/11/2017 00:00:00

Argo data format

- Official Argo data available on GDACs: **NetCDF files**
  - comply with NetCDF Climate and Forecast (CF) Metadata Conventions (version 1.6)



- If **Delayed Mode Quality Control** has been performed: « **_ADJUSTED** » fields
- Each parameter is associated with a Quality Control « FLAG »

-> **<u>All data are kept</u>**

Argo user's manual V3.4: *<u>https://doi.org/10.13155/29825</u>*

Argo Quality Control

- Real Time:

DOXY -> DOXY_QC

24 automatic tests (platform identification, impossible date, location, global range, etc…)

- Adjusted in Real-Time :

DOXY_ADJUSTED -> DOXY_ADJUSTED_QC

Automatically apply a gain correction

- Delayed Mode:

DOXY_ADJUSTED -> DOXY_ADJUSTED_QC

Time series analysis by experts
Possible correction for sensor drifts

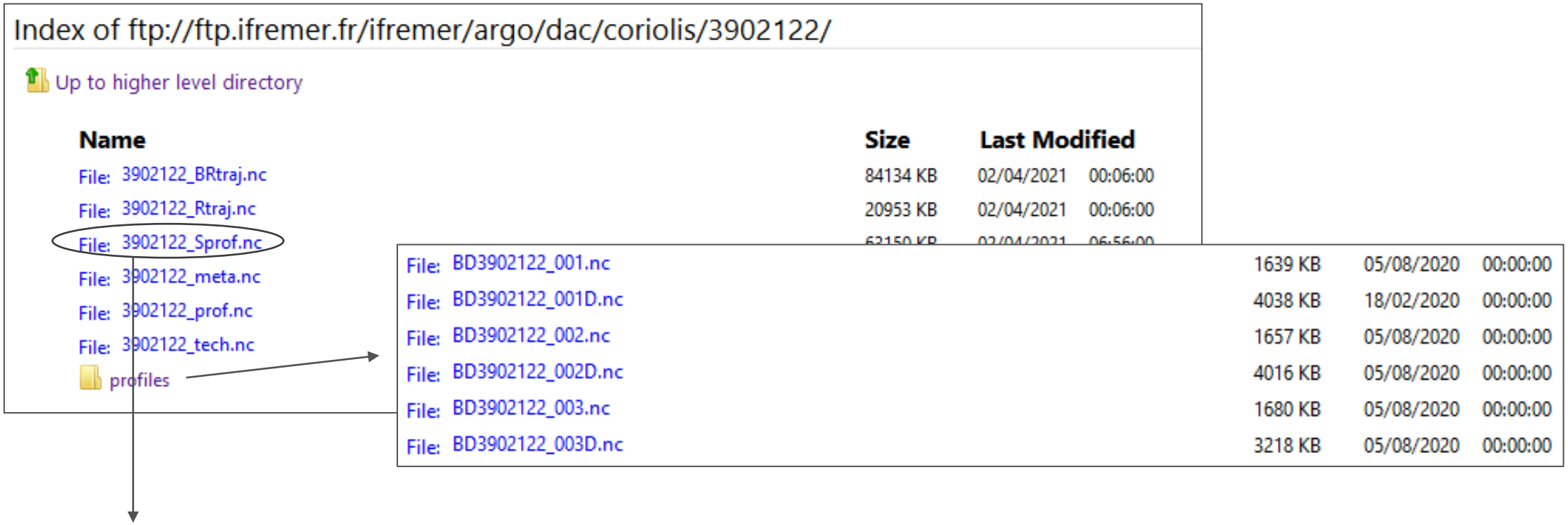
QC flag scale

0	No QC was performed
1	**Good data**
2	Probably good data
3	Bad data that are potentially correctable
4	Bad data
5	Value changed
6	Not used
7	Not used
8	Interpolated value
9	Missing value

Argo QC Manual for Argo quality control manual for biogeochemical data: http://dx.doi.org/10.13155/40879

Argo BGC data

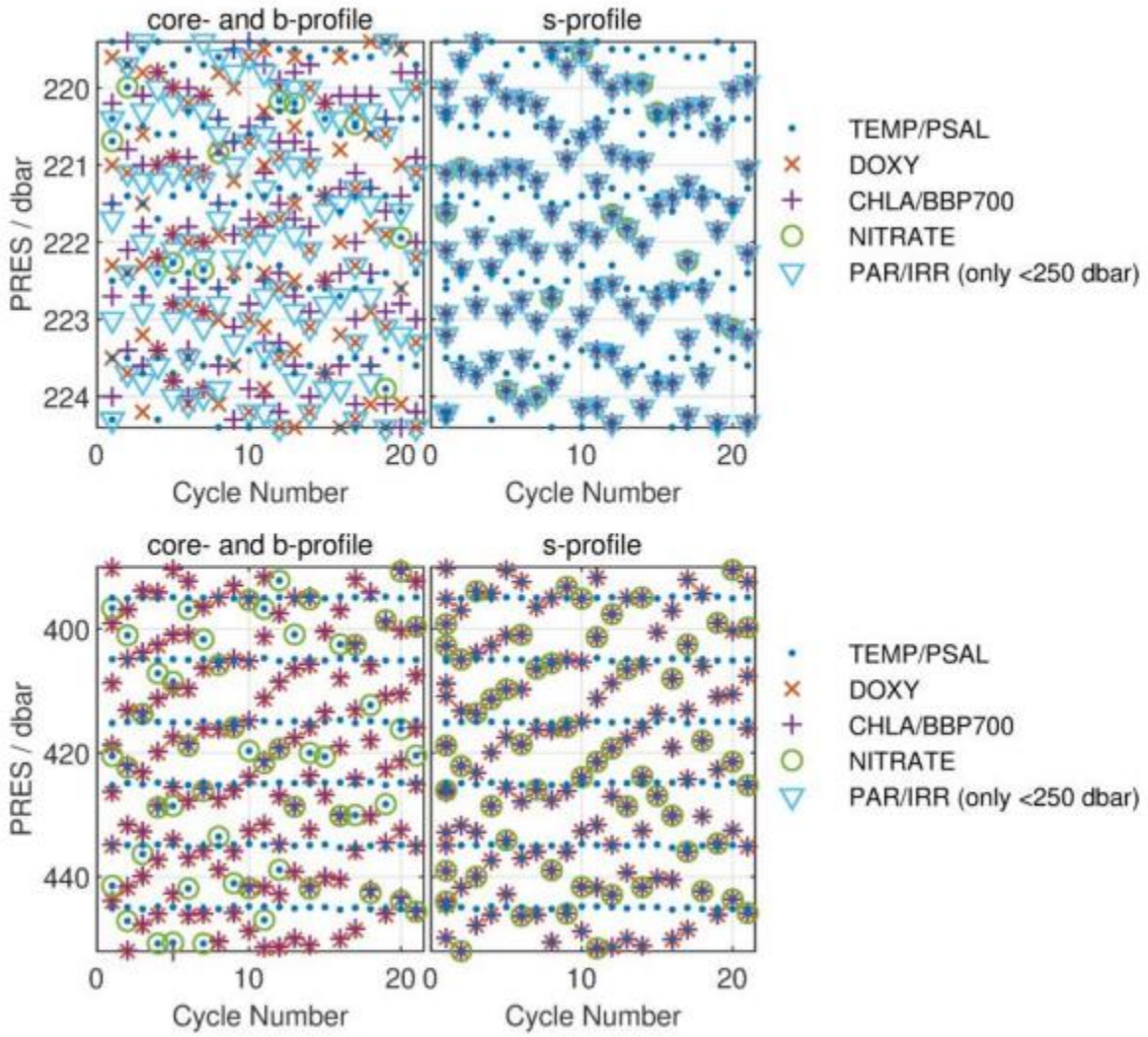
- Same philosophy as for « core Argo »



« Sprof » = synthetic profiles: all BGC variables along the same pressure axis.

Synthetic Files

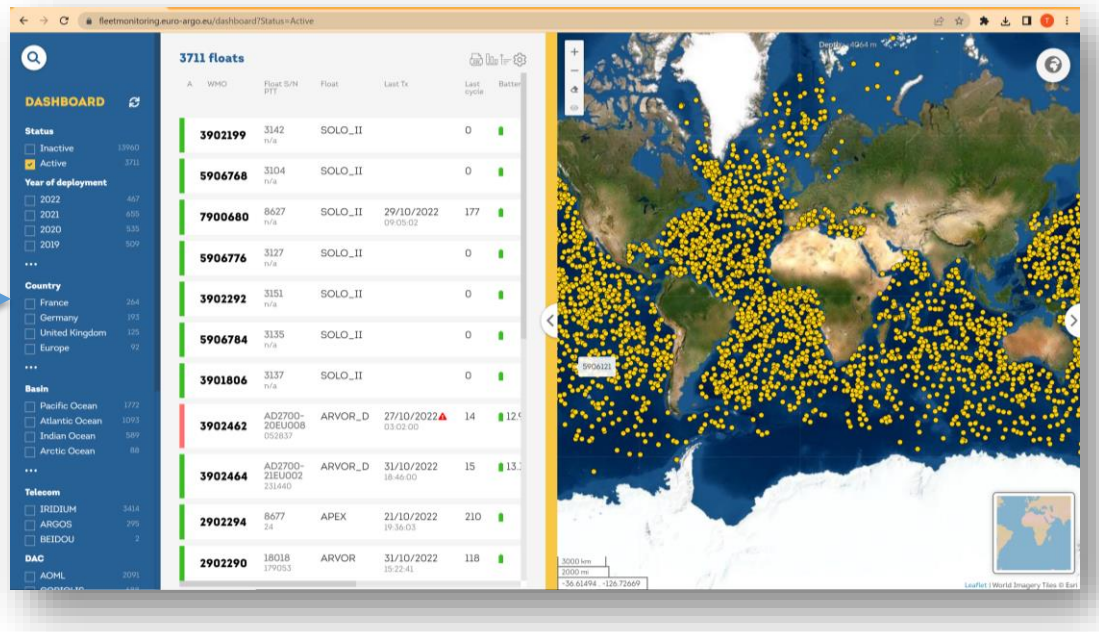
- The goal of a simplified, synthetic profile is to co-locate as many BGC observations as possible while preserving the character of the sampling pattern, i.e., sample interval, number of samples, and approximate pressure locations.



BGC-Argo synthetic profile file processing and format on Coriolis GDAC: https://doi.org/10.13155/55637

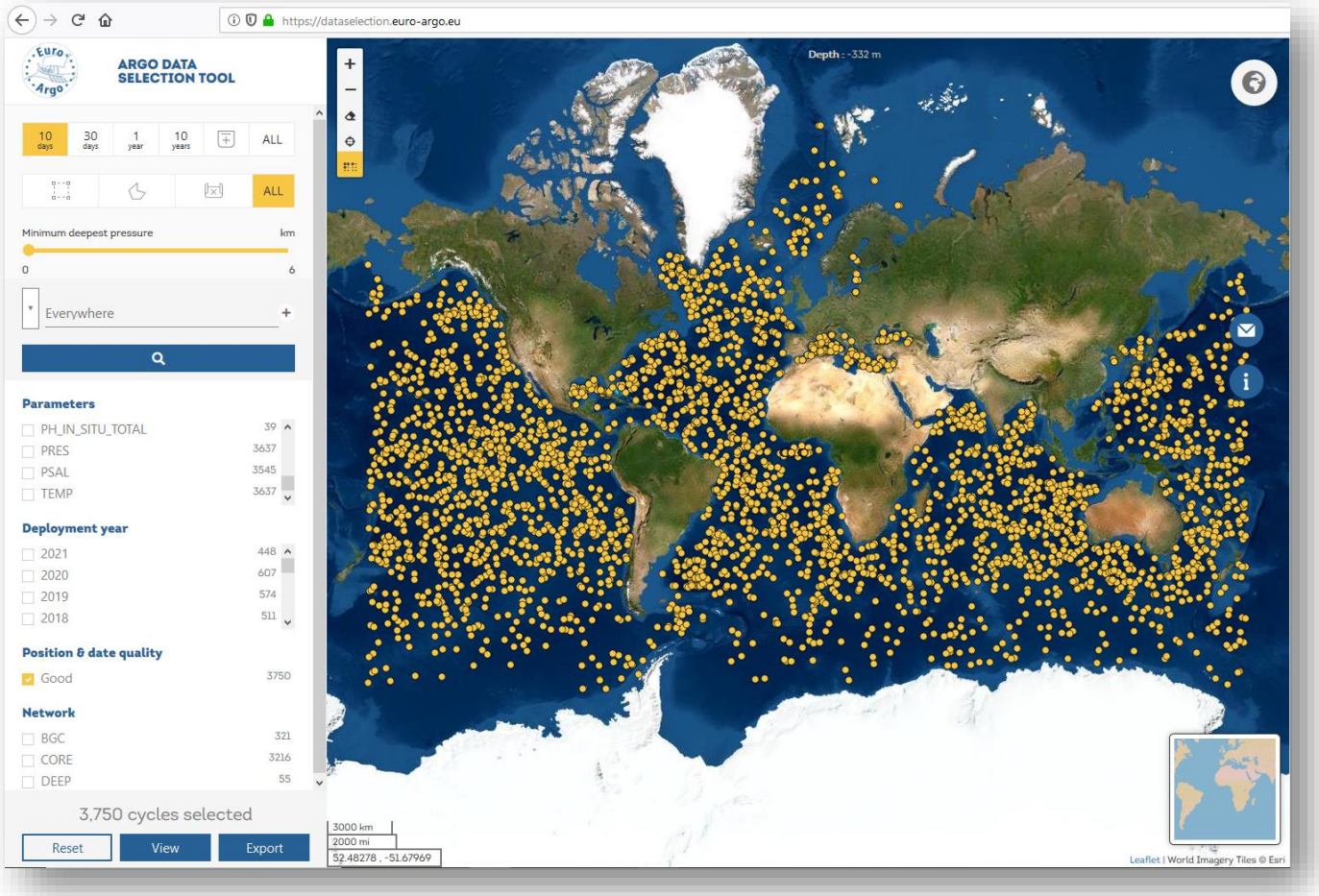
Argo data access & services

- Argo ftp server (see previous slides)
  ftp://ftp.ifremer.fr/ifremer/argo
- Argo DOIs (Data Object Identifiers)
  http://www.argodatamgt.org/Access-to-data/Argo-DOI-Digital-Object-Identifier
- Argo synchronization service (rsync)
  http://www.argodatamgt.org/Access-to-data/Argo-GDAC-synchronization-service
- Argo Thredds servers
  http://tds0.ifremer.fr/thredds/catalog/CORIOLIS-ARGO-GDAC-OBS/catalog.html
- Argo ERDDAP data server
  http://www.ifremer.fr/erddap
- Argo interactive data selection
  https://dataselection.euro-argo.eu
- Argo floats dashboard
  https://fleetmonitoring.euro-argo.eu/dashboard
- Argopy library https://github.com/euroargodev/argopy



Argo data selection tool

- https://dataselection.euro-argo.eu/
  - Enable profile selection based on several filters:
    - ✓ Spatial (either free polygone selection or by pre-defined region)
    - ✓ Temporal
    - ✓ by parameter
    - ✓ by QC
    - ✓ …
  - Export:
    - ✓ CMEMS (OceanSITE) NetCDF files
    - ✓ ASCI files
    - ✓ Argo Netcdf
  - Data visualisation



Documentation

- General information on Argo data: Data section of the Argo programme website: https://argo.ucsd.edu/data

- All Argo manuals: on the Argo Data Management website: http://www.argodatamgt.org/Documentation
  - Data format: Argo NetCDF V3.1 CF1.6 documented in “Argo user’s manual” http://dx.doi.org/10.13155/29825
  - Quality control Manual for biogeochemical data http://dx.doi.org/10.13155/40879

Coriolis

http://www.argodatamgt.org
support@argo.net

Coriolis

Argo