PRODUCT USER MANUAL

For Global Delayed Mode Insitu Dataset
INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b called CORA

Issue: 1.9

Contributors: T. Szekely

Approval Date: June 2019
### CHANGE RECORD

<table>
<thead>
<tr>
<th>Document version</th>
<th>Date</th>
<th>Change</th>
<th>Description of Change</th>
<th>Author</th>
<th>Validated By</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>10/06/2013</td>
<td>Creation of the document</td>
<td>GROUAZEL Antoine</td>
<td>S Pouliquen</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>16/01/2014</td>
<td>Update for CORA-GLOBAL 4.0</td>
<td>GROUAZEL Antoine</td>
<td>S Pouliquen</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>15/12/2014</td>
<td>Update for CORA 4.1</td>
<td>Tanguy Szekely</td>
<td>L. Crosnier</td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>May 1 2015</td>
<td>All</td>
<td>Change format to fit CMEMS graphical rules</td>
<td>L. Crosnier</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>Dec 10 2015</td>
<td>All</td>
<td>Change MyOcean in CMEMS Update for CMEMS V2</td>
<td>S. Pouliquen</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>update for CMEMS V2</td>
<td>Tanguy Szekely</td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td></td>
<td>All</td>
<td>Update for CORA4,2</td>
<td>Tanguy Szekely</td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>19/12/2016</td>
<td></td>
<td>Update for CORA 5.0</td>
<td>Tanguy Szekely</td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>01/03/2018</td>
<td></td>
<td>Update for CORA 5.1</td>
<td>Tanguy Szekely</td>
<td></td>
</tr>
<tr>
<td>1.9</td>
<td>04/04/2019</td>
<td></td>
<td>Update for CORA 5.2 and Easy CORA</td>
<td>Tanguy Szekely</td>
<td></td>
</tr>
</tbody>
</table>
Content

Content........................................................................................................................................... 3

1 INTRODUCTION.................................................................................................................................... 5
   1.1 Scope of this document........................................................................................................................ 5
   1.2 The CMEMS project............................................................................................................................ 5
   1.3 Short introduction to the product ......................................................................................................... 5

2 PRODUCT DESCRIPTION ..................................................................................................................... 6
   2.1 History of the product and links with INSITU MyOcean Distribution ............................................. 6
   2.2 Specificity of the last release .............................................................................................................. 6
   2.3 Content of the product and sub-products ............................................................................................ 8
   2.4 Organisation of the dataset on FTP server ........................................................................................... 15
   2.5 Additional documents about the dataset ............................................................................................ 16
   2.6 Validation checks and assessments .................................................................................................... 16

3 PRODUCT DISTRIBUTION ................................................................................................................ 18
   3.1 Which Download mechanism is available for this product? ............................................................... 18
   3.2 How to download this product? ......................................................................................................... 18
   3.3 Portal ftp structure............................................................................................................................... 18

4 FILES NOMENCLATURE AND FORMAT ..................................................................................... 20
   4.1 Nomenclature of files when downloaded through the Copernicus FTP Service ............................... 20
   4.2 File Format: Netedf ......................................................................................................................... 20
   4.3 Structure and semantic of NetCDF maps files ................................................................................... 21

5 Guidance for the users.......................................................................................................................... 29
   5.1 How to use DC_REFERENCE? ........................................................................................................... 29
   5.2 How to find a particular data type in CORA? .................................................................................... 29
   5.3 How to use Quality Flags? ............................................................................................................... 30
   5.4 How to use XBT with thermal offset and depth corrected? .............................................................. 30
   5.5 How to know what is the kind of data (instrument/platform) of a given measure? ......................... 31

6 Attachments.......................................................................................................................................... 33
   6.1 CORA 4.1 GLOBAL maps per type of data .................................................................................... 33
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>Climate Forecast (convention for NetCDF)</td>
</tr>
<tr>
<td>CORA</td>
<td>COriolis Re-Analysis</td>
</tr>
<tr>
<td>CORIOLIS</td>
<td>In situ data system for operational oceanography</td>
</tr>
<tr>
<td>DT</td>
<td>Delayed Time</td>
</tr>
<tr>
<td>EU</td>
<td>European Union</td>
</tr>
<tr>
<td>FTP</td>
<td>File Transfer Protocol</td>
</tr>
<tr>
<td>GDAC</td>
<td>Global Data Archiving Centre</td>
</tr>
<tr>
<td>GTS</td>
<td>Meteorological data exchange network</td>
</tr>
<tr>
<td>INS</td>
<td>In situ</td>
</tr>
<tr>
<td>ISAS</td>
<td>In Situ Analysis System</td>
</tr>
<tr>
<td>MFC</td>
<td>Monitoring and Forecasting Centre</td>
</tr>
<tr>
<td>NetCDF</td>
<td>Network Common Data Form</td>
</tr>
<tr>
<td>NRT</td>
<td>Near Real Time</td>
</tr>
<tr>
<td>OA</td>
<td>Objective Analysis</td>
</tr>
<tr>
<td>PUM</td>
<td>Product User Manual</td>
</tr>
<tr>
<td>RAN</td>
<td>ReANalysis</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>Research and Development</td>
</tr>
<tr>
<td>RT</td>
<td>Real Time</td>
</tr>
<tr>
<td>S</td>
<td>Sea Salinity</td>
</tr>
<tr>
<td>T</td>
<td>Sea Temperature</td>
</tr>
<tr>
<td>TAC</td>
<td>Thematic Assembly Centre</td>
</tr>
</tbody>
</table>
1  INTRODUCTION

1.1 Scope of this document

This Product User Manual describes CORA-GLOBAL-5.2 dataset and the additional dataset EasyCORA. These datasets are distributed as product INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b in CMEMS catalogue. This document explain how the datasets are build, what is the content, which data services are available to have access to the data.

1.2 The CMEMS project

The main objective of the COPERNICUS Marine and Environment Monitoring Service is to deliver and operate a rigorous, robust and sustainable Ocean Monitoring and Forecasting system to users for marine applications: maritime safety, marine resources, marine and coastal environment and climate, seasonal and weather forecasting. The CMEMS INSTAC prepares re-analysed datasets for reanalysis activities performed by the CMEMS MFCs and external users in collaboration with the SeaDataNet infrastructure and the EMODnet program for the global ocean and the European regional seas.

1.3 Short introduction to the product

Product INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b is generated by the company OceanScope and the Coriolis team (the Data Centre and the R&D team) in Brest, France. It provides global Temperature and Salinity observations. This dataset (also called CORA-GLOBAL, Cabanes et al., 2013, doi:10.5194/osd-9-1273-2012) is a reanalysis of these Real / Near Real Time insitu data in delayed-mode. The Easy CORA dataset in an extraction of the CORA-GLOBAL dataset providing only the best quality measurements with a vertical and temporal sub setting.
2 PRODUCT DESCRIPTION

CORA-GLOBAL-5.2 is a situ global temperature and salinity dataset that aggregates data from Coriolis database which is the IN SITU TAC Global component of the WP15 of CMEMS project. This dataset is updated twice a year by R&D Coriolis team. The April release corresponds to profiles dated up to June of the n-1 year. The November release corresponds to December of year n-1. The data are extracted in NetCDF Argo format at a given date. The dataset contains data from different types of instruments: mainly Argo floats, XBT, CTD and XCTD, and Moorings (see table 8). The data are stored in different files types: PF, XB, CT, OC, MO, BA, TE ... explained hereafter in the manual (§6). Measurements from CORA have been through both Coriolis validation (automatic checks, objective analysis) and a delayed-mode validation (format and range checks, climatological tests, spike checks, duplicates checks, XBT correction, ultimate objective analysis).

2.1 History of the product and links with INSITU MyOcean Distribution

The program Coriolis has been setup at Ifremer at the beginning of the 2000’s in the wake of the development of operational oceanography in France. The project was launched in order to provide ocean in situ measurements to the French operational ocean analysis and forecasting system (Mercator-Océan) and to contribute to a continuous, automatic, and permanent observation networks. The Coriolis data centre has been set up to gather, qualify and distribute data from the global ocean both in real and delayed time. The Coriolis database is a real time database as it is updated every day as new data arrive. On the contrary, the CORA dataset corresponds to an extraction of all in situ temperature and salinity data from the Coriolis database at a given time. All the data is then re-qualified.

The first release of CORA dataset was in 2007. Since CORA-GLOBAL3.4 it is integrated to MyOcean project as the Temperature and salinity reprocessed product (T&S REP) from the global region of the InSitu TAC. Data in CORA correspond to data in global Distribution Unit but with extra assessments. From the user point of view this dataset is more reliable than monthly latest or history distribution but it is less up-to-date since there is only twice release of CORA per year.

Moreover, the file format is not the official Copernicus file format because CORA main users are modellers that need a specific classification per date to ease the assimilation in models.

The CORA products are now diffused by the Copernicus Environment and Monitoring System.

2.2 Specificity of the last release

The main new features in CORA-GLOBAL-05.2 is the integration of:

- The delayed time mode validation of profiles from year 2017 and the first 6 month of 2018.
- Integration of measurements updated in the Coriolis database and covering the period 1950-2016
<table>
<thead>
<tr>
<th>Dataset version</th>
<th>Date of diffusion</th>
<th>Starting extraction date (netcdf)</th>
<th>Dataset time span</th>
<th>Severity</th>
<th>New Features</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>CORA1.0</td>
<td>2007</td>
<td>1990-2005</td>
<td>Major</td>
<td>Beta version</td>
<td>Emmanuelle Autret</td>
<td></td>
</tr>
<tr>
<td>CORA2.2</td>
<td>2009</td>
<td>1990-2008</td>
<td>Major</td>
<td>+ procedures de validation (climatological checks) + analyse objective (gridded fields)</td>
<td>Cécile Pertuisot Cécile Cabanes</td>
<td></td>
</tr>
<tr>
<td>CORA3.2</td>
<td>September 2011</td>
<td>14/03/2011</td>
<td>1990-2010</td>
<td>Minor</td>
<td>+2010 + duplicate detection/suppression + XBT correction from Hamon et al, 2011 + Sea mammals data + model assimilation feedbacks (glorys1v1)</td>
<td>Cécile Cabanes Antoine Grouazel</td>
</tr>
<tr>
<td>CORA3.3</td>
<td>September 2012</td>
<td>January to September 2011: 04/10/2011 and October to December 2011: 22/03/2012</td>
<td>1990-2011</td>
<td>Minor</td>
<td>+2011 + Article submitted to Ocean Science + integration of alerts from objective analysis standardisation</td>
<td>Antoine Grouazel</td>
</tr>
<tr>
<td>CORA3.4</td>
<td>April 2013</td>
<td>22/01/2013</td>
<td>1990-2011</td>
<td>Minor</td>
<td>+ ICES data from coriolis database for V3 of MyOcean.</td>
<td>Antoine Grouazel</td>
</tr>
</tbody>
</table>
| **CORA4.0** | January 2014 (V4 Myocean) | 28/05/2013 up to 15/06/2013 | 1990-2012 | Major | +2012  
+Time series data (TSG, FerryBox, Drifters, Moorings)  
+Integration of alerts from objective analysis residuals | Antoine Grouazel  
Emilie Brion |
| **CORA 4.1** | January 2015 | 02/01/2015 | 1950-1989 (Extraction from EN4) + 1990-2013 | Minor | + 2013  
+ TSG data, delayed mode sea mammals and surface drifters  
+ Integration of XBT and CTD from EN4 | Tanguy Szekely  
Emilie Brion  
Jerôme Gourrion |
| **CORA 4.2** | January 2015 | 15/01/2015 | 1950-1989 (Extraction from EN4) + 1990-2014 | Major | + 2014  
+ 2,4 million profiles from the SHOM (Service Hydrographique de la Marine), mostly of XBTs CTDs and MBTs from 1950 to 2009 | Tanguy Szekely  
Emilie Brion  
Jerôme Gourrion |
| **CORA5.0** | April 2017 | 15/05/2016 | 1950-2015 | Major | + 2015  
+ profiles extracted from the EN4 dataset and validated | Tanguy Szekely  
Anne Piron  
Jerôme Gourrion |
| **CORA 5.1** | April 2018 | 15/05/2017 | 1950-2016 | Minor | Validation in delayed time mode of year 2016 and of the profiles updated from the Coriolis database since the last CORA version. | Tanguy Szekely  
Anne Piron  
Jerôme Gourrion |
| **CORA 5.2** | April 2019 | 15/05/2018 | 1950- June 2018 | Major | + Validation in delayed time mode of year 2017 and jan-june 2018 and of the profiles updated from the Coriolis database since the last CORA version.  
+ Production of the first version of the Easy CORA dataset | Tanguy Szekely  
Anne Piron  
Jerôme Gourrion  
Nathalie Verbruge  
Rachel Killick |

*Table 1: CORA releases*
2.3 Content of the product and sub-products

The Coriolis centre (global region from the InSitu TAC) receives data from Argo GDAC, French research ships, GTS flow, GTSP, GOSUD GDAC, MEDS, voluntary observing and merchants ships, moorings (TAO-TRITON-PIRATA-RAMA plus coastal moorings), and the World Ocean Database (not in real time for the last one and for CTD only). CORA thus contains data from different types of instruments: mainly Argo floats, XBT, CTD, XCTD, and moorings. Figure 1 gives an overview of the yearly number of CORA profiles and the evolution of instruments types. Table 4 gives an overview of the data providers for delayed time mode and near real time mode data for various instruments types.

Figure 1: Yearly number of profiles in CORA dataset for various instrument types.

<table>
<thead>
<tr>
<th>Type of measurements</th>
<th>Real time Sources</th>
<th>Delayed Mode Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gliders</td>
<td>Most EGO gliders data</td>
<td></td>
</tr>
<tr>
<td>Drifters</td>
<td>GTSP</td>
<td>MétéoFrance current reanalysis on SVP buoys, UDASH database</td>
</tr>
<tr>
<td>Moorings</td>
<td>TAO-TRITON-PIRATA-RAMA + others moorings from OceanSites network</td>
<td>PMEL reprocessed files, UDASH database</td>
</tr>
<tr>
<td>Profilers</td>
<td>Argo GDAC</td>
<td>Argo GDAC with PI adjustments, UDASH database</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Thermo-salinographs</td>
<td>SSS merchant and research ships</td>
<td>GOSUD GDAC, UDASH database</td>
</tr>
<tr>
<td>Sea mammals</td>
<td>Data from CEBC animals trough GTSPP</td>
<td>MEMO high resolution measurements, UDASH database</td>
</tr>
<tr>
<td>XBTs, CTDs, MBTs, X-CTD, bottles, scan-fish</td>
<td>GTSPP (bathy or tessac)+ French research ships, Copernicus regional INS TAC</td>
<td>ICES, SeaDatanet, WOD, UDASH database, ...</td>
</tr>
</tbody>
</table>

*Table 2: Coriolis data sources*
2.5 Additional documents about the dataset

The following manual applies to the following list of products described in Copernicus Catalog

<table>
<thead>
<tr>
<th>Short Description</th>
<th>Product code</th>
<th>Area</th>
<th>Delivery Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global DT</td>
<td>INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b (also called CORA-GLOBAL)</td>
<td>Global</td>
<td>Yearly</td>
</tr>
</tbody>
</table>

Table 3: List of INS TAC products for which this manual apply

Detailed information on the INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b product is available at:

- **If you use the CORA data, please cite this article that describe CORA dataset:**


2.6 Validation checks and assessments

To have a detailed description of the tests performed on CORA dataset please refer to the CORA product QUID document distributed in the Copernicus catalogue.

Following paragraph gives a short summary of those assessments.

Data extracted from Coriolis database has undergoes automatic checks and objectives analysis within the database. After the extraction it goes through a series of extra validation in delayed-mode: this is the added value of CORA dataset.

<table>
<thead>
<tr>
<th>Name of the validation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure on earth</td>
<td>Compare position to bathymetry, to reject bad positions</td>
</tr>
<tr>
<td>Date check</td>
<td>Checks that the date correspond to the name of the file</td>
</tr>
<tr>
<td>Parameter Range check</td>
<td>Check that TEMP PSAL PRES and DEPH have acceptable values</td>
</tr>
<tr>
<td>Constant check</td>
<td>Check that TEMP PSAL PRES and DEPH have different values along the vertical</td>
</tr>
<tr>
<td>Ascending immersion check</td>
<td>Check that PRES and/or DEPH are increasing</td>
</tr>
<tr>
<td>Duplicate levels check</td>
<td>Check that immersion level are not duplicated</td>
</tr>
<tr>
<td>Minmax test</td>
<td>Check the profiles by comparing the measurements with the minimum and maximum temperature/salinity ever measured with the ARGO floats from 2002 to 2014.</td>
</tr>
<tr>
<td>Name of the validation</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Spike check</td>
<td>Semi-automatic (need a visual control) check that spot spikes on profiles</td>
</tr>
<tr>
<td>Quality Flag relevance</td>
<td>Control that a given QC is relevant with the associated measure</td>
</tr>
<tr>
<td>Depth wrote in pressure field</td>
<td>Check that depth measurements are not written in PRES field (especially for XBT and X-CTD)</td>
</tr>
<tr>
<td>Duplicate profile</td>
<td>Detect duplicate profile and delete the one having a worse quality or less meta-data</td>
</tr>
<tr>
<td>XBT correction</td>
<td>Empirical correction of depth bias and temperature offset on XBTs. Method of Hamon et al based on co-localisation with CTD profiles.</td>
</tr>
<tr>
<td>Assimilation feedback</td>
<td>Alerts on profiles raised by a too strong innovation value when assimilated in a model.</td>
</tr>
<tr>
<td>Ultimate objective analysis</td>
<td>Last step of the process that guarantee a global (spatial and temporal) consistency of the dataset by computing a residual value between each measure and the “background” given by the climatology and the other measurements.</td>
</tr>
</tbody>
</table>

*Table 4: CORA Validation process*
3.1 Which Download mechanism is available for this product?

The data files can be downloaded from the Copernicus Marine service catalogue, via a ftp protocol: http://marine.copernicus.eu/services-portfolio/access-to-products/

3.2 How to download this product?

You first need to register. Please find the registration steps on our website:

http://marine.copernicus.eu/web/34-products-and-services-faq.php#1

Once registered, the CMEMS FAQ http://marine.copernicus.eu/web/34-products-and-services-faq.php#1 will guide you on how to download a product through the CMEMS Web Portal Subsetter and FTP Services.

3.3 Portal ftp structure

The original product, containing temperature and salinity measure at observed levels, is called INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b.

To give more visibility to the gridded fields and standardized data that are a derived product of INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b, a new product has been added to the catalogue called INSITU_GLO_TS_OA_REP_OBSERVATIONS_013_002_b.

Those by-products are systemically built at the end of the validation process of the original data set thanks to the software ISAS dedicated to run objective analysis.

The ftp portal described here correspond to the current version of the dataset (CORA 5.2). To have a detailed description of the previous version of the dataset, please see the corresponding version of this document.

To recap CMEMS INSTAC service distributes on its FTP server:

- “RAW” data that have been through the validation process in delayed-mode: (RAW/) INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b. Those raw data are distributed at global scale and also in splited version for each CMEMS region.

- global gridded fields (77°S,89.89°N,180°W,180°E,z=0-2000m, 720x545x152 : longitude latitude level) build at the end of the validation process: (OA/FIELD/) INSITU_GLO_TS_OA_REP_OBSERVATIONS_013_002_b

- geographically and daily sub-sampled data interpolated on 152 vertical levels: (OA/DATA/) INSITU_GLO_TS_OA_REP_013_030_b

- yearly Index files to browse quickly meta-data from the dataset: (Index/)
The global datasets and the corresponding extractions are distributed in the ftp portal.

The following directories of the ftp portal relate to various and different data:

- CORIOLIS-GLOBAL-CORA5.2: root directory
- CORA-GLOBAL: Directory containing the CORA dataset.
- EasyCORA: Directory containing the EasyCORA dataset.

Inside the CORA-GLOBAL and EasyCORA directory, the datasets are distributed in directory corresponding to the CMEMS sub region: arctic, Baltic, Barrens, black sea, mead sea, northwestshelf, southwestshelf and global.

For detailed information on CORA directories, please visit:

4.1 Nomenclature of files when downloaded through the Copernicus FTP Service

INSITU_GLO_TSREP_OBSERVATIONS_013_001_b files nomenclature when downloaded through the Copernicus FTP service is based as follows:

- The INSITU files: CO_code/YYYYMMDDD.PR_TT.nc where:
  - code is the name of the analysis performed: DMQCGL01
  - YYYYMMDDD is the date of the data,
  - PR stands for vertical Profile, TS stands for time serie
  - TT is the type of file (see 15)
CORA file types content:
The CORA file system is based on 17 daily file types corresponding to the instrument type of the data provider.

- **PR_PF files**: data from Argo floats directly received from DACS (real Time and delayed mode if available). These data have a nominal accuracy of 0.01° and 0.01 PSU and are transmitted with full resolution.
- **PR_XB files**: XBT or XCTD data received from research and opportunity vessels have accuracy within 0.03° to 0.1° for temperature and 0.03 to 0.1 PSU for salinity.
- **PR_CT files**: contains CTD data from research vessels (accuracy on the order of 0.002° for temperature and 0.003 PSU for salinity after calibration) but also data from sea mammals equipped with CTD (accuracy is on the order of 0.01° for temperature and 0.02 PSU for salinity but can be lower depending on the availability of reference data for post-processing, see Boehme et al, 2009) and received from MNHN and some sea Gliders.
- **PR_OC files**: Others CTD and XCTD data coming from the high resolution CTD dataset of the World ocean database 2009 (WOD09).
- **PR_TE and BA files**: The two last categories are for all the data transmitted through the GTS (data from Argo floats not yet received at the DACS, mooring, XBT,...). This transmission system imposes limitation on the accuracy: data is truncated two and one places beyond decimal point for TE and BA type respectively.
- **PR_IC files**: CTD from ICES dataset gathered by Danish CIUEM. Those profiles complete the CTDs coverage on the period 1990-2011.
- **PR_SH files**: About 2,6 milions profiles from the SHOM database. Most of them cover the 1950-1990 period.
- **PR_SM files**: Profiles measured by ser mammals mounted sensors.
- **PR_ME files**: CTD from SISMER database, coming from French oceanographic campaigns. First loading in GLOBAL INS TAC in July 2013 (about 60.000 CTDs)
- **PR_OS files**: OceanSites data are mostly CTD, moorings data from Oceansites can be found in TS_MO.
- **PR_UD files**: Arctic ocean profiles extracted from the UDASH database. [https://doi.org/10.1594/PANGAEA.872931](https://doi.org/10.1594/PANGAEA.872931).
- **TS_TS files**: thermosalinograph from research ships ORE SSS (Observatory of research in Environment - Sea Surface Salinity) validated in delayed mode by scientists at US IMAGO.
- **TS_FB files**: ferry-boxes data.
- **TS_DB files**: time series from drifter buoys (mainly SVP model).
- **TS_MO files**: moorings time series containing most of the moorings data. For example, TAO TRITON RAMA and PIRATA.

### 4.2 File Format: Netcdf

The products are stored using the NetCDF format (NetCDF 3.4).

NetCDF (network Common Data Form) is an interface for array-oriented data access and a library that provides an implementation of the interface. The netCDF library also defines a machine-independent format for representing scientific data. Together, the interface, library, and format support the creation, access, and sharing of scientific data. The netCDF software was developed at the Unidata Program Center in Boulder, Colorado. The netCDF libraries define a machine-independent format for representing scientific data.

Please see Unidata netCDF pages for more information, and to retrieve netCDF software package.

---

1 Principal Investigators have to provide a delayed mode validation on their floats within 6 months after the date of observation.
NetCDF data is:

- Self-Describing. A netCDF file includes information about the data it contains.
- Architecture-independent. A netCDF file is represented in a form that can be accessed by computers with different ways of storing integers, characters, and floating-point numbers.
- Direct-access. A small subset of a large dataset may be accessed efficiently, without first reading through all the preceding data.
- Append-able. Data can be appended to a netCDF dataset along one dimension without copying the dataset or redefining its structure. The structure of a netCDF dataset can be changed, though this sometimes causes the dataset to be copied.
- Shareable. One writer and multiple readers may simultaneously access the same netCDF file.

For INSITU files, the used NetCDF format is Argo 3.3 (see http://www.argodatamgt.org/Documentation/Argo-NetCDF-sample-files) and for other files it is CF-1.4.
4.3 Structure and semantic of NetCDF maps files

For CO_code_YYYYMMDD_PR_TT.nc:

```plaintext
netcdf CO_RTOAGL01_20130101_PR_BA {

dimensions:

    DATE_TIME = 14 ;
    STRING256 = 256 ;
    STRING64 = 64 ;
    STRING32 = 32 ;
    STRING16 = 16 ;
    STRING8 = 8 ;
    STRING4 = 4 ;
    STRING2 = 2 ;
    N_PROF = 49 ;
    N_PARAM = 2 ;
    N_LEVELS = 473 ;
    N_CALIB = 1 ;
    N_HISTORY = UNLIMITED ; // (0 currently)

variables:

    char DATA_TYPE(STRING16);
    DATA_TYPE:long_name = "Data type";
    DATA_TYPE:_FillValue = " " ;
    char FORMAT_VERSION(STRING4);
    FORMAT_VERSION:long_name = "File format version";
    FORMAT_VERSION:_FillValue = " " ;
    char HANDBOOK_VERSION(STRING4);
    HANDBOOK_VERSION:long_name = "Data handbook version";
    HANDBOOK_VERSION:_FillValue = " " ;
    char REFERENCE_DATE_TIME(DATE_TIME);
    REFERENCE_DATE_TIME:long_name = "Date of reference for Julian days";
    REFERENCE_DATE_TIME:conventions = "YYYYMMDDHHMISS" ;
    REFERENCE_DATE_TIME:_FillValue = " " ;
    char DATE_CREATION(DATE_TIME);
    DATE_CREATION:long_name = "Date of file creation";
    DATE_CREATION:conventions = "YYYYMMDDHHMISS" ;
    DATE_CREATION:_FillValue = " " ;
    char DATE_UPDATE(DATE_TIME);
    DATE_UPDATE:long_name = "Date of update of this file";
```
DATE_UPDATE:conventions = "YYYYMMDDHHMISS";
DATE_UPDATE:_FillValue = " ";
char PLATFORM_NUMBER(N_PROF, STRING8);
PLATFORM_NUMBER:long_name = "Float unique identifier";
PLATFORM_NUMBER:conventions = "WMO float identifier : A9IIII";
PLATFORM_NUMBER:_FillValue = " ";
char PROJECT_NAME(N_PROF, STRING64);
PROJECT_NAME:long_name = "Name of the project";
PROJECT_NAME:_FillValue = " ";
char PI_NAME(N_PROF, STRING64);
PI_NAME:long_name = "Name of the principal investigator";
PI_NAME:_FillValue = " ";
char STATION_PARAMETERS(N_PROF, N_PARAM, STRING16);
STATION_PARAMETERS:long_name = "List of available parameters for the station";
STATION_PARAMETERS:conventions = "Argo reference table 3";
STATION_PARAMETERS:_FillValue = " ";
int CYCLE_NUMBER(N_PROF);
CYCLE_NUMBER:long_name = "Float cycle number";
CYCLE_NUMBER:conventions = "0..N, 0 : launch cycle (if exists), 1 : first complete cycle";
CYCLE_NUMBER:_FillValue = 99999;
char DIRECTION(N_PROF);
DIRECTION:long_name = "Direction of the station profiles";
DIRECTION:conventions = "A: ascending profiles, D: descending profiles";
DIRECTION:_FillValue = " ";
char DATA_CENTRE(N_PROF, STRING2);
DATA_CENTRE:long_name = "Data centre in charge of float data processing";
DATA_CENTRE:conventions = "Argo reference table 4";
DATA_CENTRE:_FillValue = " ";
char DC_REFERENCE(N_PROF, STRING32);
DC_REFERENCE:long_name = "Station unique identifier in data centre";
DC_REFERENCE:conventions = "Data centre convention";
DC_REFERENCE:_FillValue = " ";
char DATA_STATE_INDICATOR(N_PROF, STRING4);
DATA_STATE_INDICATOR:long_name = "Degree of processing the data have passed through";
DATA_STATE_INDICATOR:conventions = "Argo reference table 6";
DATA_STATE_INDICATOR:_FillValue = " ";
char DATA_MODE(N_PROF);
DATA_MODE:long_name = "Delayed mode or real time data";
DATA_MODE:conventions = "R : real time; D : delayed mode; A : real time with adjustment";
DATA_MODE:_FillValue = " ";

char INST_REFERENCE(N_PROF, STRING64);
    INST_REFERENCE:long_name = "Instrument type";
    INST_REFERENCE:conventions = "Brand, type, serial number";
    INST_REFERENCE:_FillValue = " ";

char FIRMWARE_VERSION(N_PROF, STRING16);
    FIRMWARE_VERSION:long_name = "Instrument version";
    FIRMWARE_VERSION:conventions = "";
    FIRMWARE_VERSION:_FillValue = " ";

char WMO_INST_TYPE(N_PROF, STRING4);
    WMO_INST_TYPE:long_name = "Coded instrument type";
    WMO_INST_TYPE:conventions = "Argo reference table 8";
    WMO_INST_TYPE:_FillValue = " ";

double JULD(N_PROF);
    JULD:long_name = "Julian day (UTC) of the station relative to REFERENCE_DATE_TIME";
    JULD:standard_name = "time";
    JULD:units = "days since 1950-01-01 00:00:00 UTC";
    JULD:conventions = "Relative julian days with decimal part (as parts of day)";
    JULD:_FillValue = 999999.0;
    JULD:axis = "T";

char JULD_QC(N_PROF);
    JULD_QC:long_name = "Quality on Date and Time";
    JULD_QC:conventions = "Argo reference table 2";
    JULD_QC:_FillValue = " ";

double JULD_LOCATION(N_PROF);
    JULD_LOCATION:long_name = "Julian day (UTC) of the location relative to REFERENCE_DATE_TIME";
    JULD_LOCATION:units = "days since 1950-01-01 00:00:00 UTC";
    JULD_LOCATION:conventions = "Relative julian days with decimal part (as parts of day)";
    JULD_LOCATION:_FillValue = 999999.0;

double LATITUDE(N_PROF);
    LATITUDE:long_name = "Latitude of the station, best estimate";
    LATITUDE:standard_name = "latitude";
    LATITUDE:units = "degree_north";
    LATITUDE:_FillValue = 999999.0;
    LATITUDE:valid_min = -90.0;
    LATITUDE:valid_max = 90.0;
    LATITUDE:axis = "Y";

double LONGITUDE(N_PROF);
    LONGITUDE:long_name = "Longitude of the station, best estimate";
LONGITUDE:standard_name = "longitude";
LONGITUDE:units = "degree_east";
LONGITUDE:_FillValue = 99999. ;
LONGITUDE:valid_min = -180. ;
LONGITUDE:valid_max = 180. ;
LONGITUDE:axis = "X";
char POSITION_QC(N_PROF);
    POSITION_QC:long_name = "Quality on position (latitude and longitude)";
    POSITION_QC:conventions = "Argo reference table 2";
    POSITION_QC:_FillValue = " ";
char POSITIONING_SYSTEM(N_PROF, STRING8);
    POSITIONING_SYSTEM:long_name = "Positioning system";
    POSITIONING_SYSTEM:_FillValue = " ";
char PROFILE_DEPH_QC(N_PROF);
    PROFILE_DEPH_QC:long_name = "Global quality flag of DEPH profile";
    PROFILE_DEPH_QC:conventions = "Argo reference table 2a";
    PROFILE_DEPH_QC:_FillValue = " ";
char PROFILE_TEMP_QC(N_PROF);
    PROFILE_TEMP_QC:long_name = "Global quality flag of TEMP profile";
    PROFILE_TEMP_QC:conventions = "Argo reference table 2a";
    PROFILE_TEMP_QC:_FillValue = " ";
char VERTICAL_SAMPLING_SCHEME(N_PROF, STRING256);
    VERTICAL_SAMPLING_SCHEME:long_name = "Vertical sampling scheme";
    VERTICAL_SAMPLING_SCHEME:conventions = "Argo reference table 16";
    VERTICAL_SAMPLING_SCHEME:_FillValue = " ";
float DEPH(N_PROF, N_LEVELS);
    DEPH:long_name = "DEPTH UNDER SEA SURFACE";
    DEPH:standard_name = "depth";
    DEPH:_FillValue = 99999.f ;
    DEPH:units = "meter";
    DEPH:valid_min = 0.f ;
    DEPH:valid_max = 12000.f ;
    DEPH:C_format = "%7.1f";
    DEPH:FORTRAN_format = "F7.1";
    DEPH:resolution = 0.1f ;
char DEPH_QC(N_PROF, N_LEVELS);
    DEPH_QC:long_name = "quality flag";
    DEPH_QC:conventions = "Argo reference table 2";
    DEPH_QC:_FillValue = " " ;
float DEPH_ADJUSTED(N_PROF, N_LEVELS);
    DEPH_ADJUSTED:long_name = "DEPTH UNDER SEA SURFACE";
    DEPH_ADJUSTED:_FillValue = 99999.f;
    DEPH_ADJUSTED:units = "meter";
    DEPH_ADJUSTED:valid_min = 0.f;
    DEPH_ADJUSTED:valid_max = 12000.f;
    DEPH_ADJUSTED:C_format = "%7.1f";
    DEPH_ADJUSTED:FORTRAN_format = "F7.1";
    DEPH_ADJUSTED:resolution = 0.1f;
char DEPH_ADJUSTED_QC(N_PROF, N_LEVELS);
    DEPH_ADJUSTED_QC:long_name = "quality flag";
    DEPH_ADJUSTED_QC:conventions = "Argo reference table 2";
    DEPH_ADJUSTED_QC:_FillValue = " ";
float DEPH_ADJUSTED_ERROR(N_PROF, N_LEVELS);
    DEPH_ADJUSTED_ERROR:long_name = "DEPTH UNDER SEA SURFACE";
    DEPH_ADJUSTED_ERROR:_FillValue = 99999.f;
    DEPH_ADJUSTED_ERROR:units = "meter";
    DEPH_ADJUSTED_ERROR:valid_min = 0.f;
    DEPH_ADJUSTED_ERROR:valid_max = 12000.f;
    DEPH_ADJUSTED_ERROR:C_format = "%7.1f";
    DEPH_ADJUSTED_ERROR:FORTRAN_format = "F7.1";
    DEPH_ADJUSTED_ERROR:resolution = 0.1f;
float TEMP(N_PROF, N_LEVELS);
    TEMP:long_name = "SEA TEMPERATURE IN SITU ITS-90 SCALE";
    TEMP:standard_name = "sea_water_temperature";
    TEMP:_FillValue = 99999.f;
    TEMP:units = "degree_Celsius";
    TEMP:valid_min = -2.f;
    TEMP:valid_max = 40.f;
    TEMP:C_format = "%9.3f";
    TEMP:FORTRAN_format = "F9.3";
    TEMP:resolution = 0.001f;
char TEMP_QC(N_PROF, N_LEVELS);
    TEMP_QC:long_name = "quality flag";
    TEMP_QC:conventions = "Argo reference table 2";
    TEMP_QC:_FillValue = " ";
float TEMP_ADJUSTED(N_PROF, N_LEVELS);
    TEMP_ADJUSTED:long_name = "SEA TEMPERATURE IN SITU ITS-90 SCALE";
    TEMP_ADJUSTED:_FillValue = 99999.f;
TEMP_ADJUSTED:units = "degree_Celsius";
TEMP_ADJUSTED:valid_min = -2.f;
TEMP_ADJUSTED:valid_max = 40.f;
TEMP_ADJUSTED:C_format = "%%9.3f";
TEMP_ADJUSTED:FORTRAN_format = "F9.3";
TEMP_ADJUSTED:resolution = 0.001f;
char TEMP_ADJUSTED_QC(N_PROF, N_LEVELS) ;
TEMP_ADJUSTED_QC:long_name = "quality flag";
TEMP_ADJUSTED_QC:conventions = "Argo reference table 2";
TEMP_ADJUSTED_QC:_FillValue = " " ;
float TEMP_ADJUSTED_ERROR(N_PROF, N_LEVELS) ;
TEMP_ADJUSTED_ERROR:long_name = "SEA TEMPERATURE IN SITU ITS-90 SCALE";
TEMP_ADJUSTED_ERROR:_FillValue = 99999.f;
TEMP_ADJUSTED_ERROR:units = "degree_Celsius";
TEMP_ADJUSTED_ERROR:C_format = "%%9.3f";
TEMP_ADJUSTED_ERROR:FORTRAN_format = "F9.3";
TEMP_ADJUSTED_ERROR:resolution = 0.001f;
char PARAMETER(N_PROF, N_CALIB, N_PARAM, STRING16) ;
PARAMETER:long_name = "List of parameters with calibration information";
PARAMETER:conventions = "Argo reference table 3";
PARAMETER:_FillValue = " " ;
char SCIENTIFIC_CALIB_EQUATION(N_PROF, N_CALIB, N_PARAM, STRING256) ;
SCIENTIFIC_CALIB_EQUATION:long_name = "Calibration equation for this parameter";
SCIENTIFIC_CALIB_EQUATION:_FillValue = " " ;
char SCIENTIFIC_CALIB_COEFFICIENT(N_PROF, N_CALIB, N_PARAM, STRING256) ;
SCIENTIFIC_CALIB_COEFFICIENT:long_name = "Calibration coefficients for this equation";
SCIENTIFIC_CALIB_COEFFICIENT:_FillValue = " " ;
char SCIENTIFIC_CALIB_COMMENT(N_PROF, N_CALIB, N_PARAM, STRING256) ;
SCIENTIFIC_CALIB_COMMENT:long_name = "Comment applying to this parameter calibration";
SCIENTIFIC_CALIB_COMMENT:_FillValue = " " ;
char SCIENTIFIC_CALIB_DATE(N_PROF, N_CALIB, N_PARAM, DATE_TIME) ;
SCIENTIFIC_CALIB_DATE:long_name = "Date of calibration";
SCIENTIFIC_CALIB_DATE:_FillValue = " " ;
char HISTORY_INSTITUTION(N_HISTORY, N_PROF, STRING4) ;
HISTORY_INSTITUTION:long_name = "Institution which performed action";
HISTORY_INSTITUTION:conventions = "Argo reference table 4";
HISTORY_INSTITUTION:_FillValue = " " ;
char HISTORY_STEP(N_HISTORY, N_PROF, STRING4) ;
HISTORY_STEP:long_name = "Step in data processing";
HISTORY_STEP:conventions = "Argo reference table 12";
HISTORY_STEP:_FillValue = ";
char HISTORY_SOFTWARE(N_HISTORY, N_PROF, STRING4);
HISTORY_SOFTWARE:long_name = "Name of software which performed action";
HISTORY_SOFTWARE:conventions = "Institution dependent";
HISTORY_SOFTWARE:_FillValue = ";
char HISTORY_SOFTWARE_RELEASE(N_HISTORY, N_PROF, STRING4);
HISTORY_SOFTWARE_RELEASE:long_name = "Version/release of software which performed action";
HISTORY_SOFTWARE_RELEASE:conventions = "Institution dependent";
HISTORY_SOFTWARE_RELEASE:_FillValue = ";
char HISTORY_REFERENCE(N_HISTORY, N_PROF, STRING64);
HISTORY_REFERENCE:long_name = "Reference of database";
HISTORY_REFERENCE:conventions = "Institution dependent";
HISTORY_REFERENCE:_FillValue = ";
char HISTORY_DATE(N_HISTORY, N_PROF, DATE_TIME);
HISTORY_DATE:long_name = "Date the history record was created";
HISTORY_DATE:conventions = "YYYYMMDDHHMISS";
HISTORY_DATE:_FillValue = ";
char HISTORY_ACTION(N_HISTORY, N_PROF, STRING4);
HISTORY_ACTION:long_name = "Action performed on data";
HISTORY_ACTION:conventions = "Argo reference table 7";
HISTORY_ACTION:_FillValue = ";
char HISTORY_PARAMETER(N_HISTORY, N_PROF, STRING16);
HISTORY_PARAMETER:long_name = "Station parameter action is performed on";
HISTORY_PARAMETER:conventions = "Argo reference table 3";
HISTORY_PARAMETER:_FillValue = ";
float HISTORY_START_PRES(N_HISTORY, N_PROF);
HISTORY_START_PRES:long_name = "Start pressure action applied on";
HISTORY_START_PRES:_FillValue = 99999.f;
HISTORY_START_PRES:units = "decibar";
float HISTORY_STOP_PRES(N_HISTORY, N_PROF);
HISTORY_STOP_PRES:long_name = "Stop pressure action applied on";
HISTORY_STOP_PRES:_FillValue = 99999.f;
HISTORY_STOP_PRES:units = "decibar";
float HISTORY_PREVIOUS_VALUE(N_HISTORY, N_PROF);
HISTORY_PREVIOUS_VALUE:long_name = "Parameter/Flag previous value before action";
HISTORY_PREVIOUS_VALUE:_FillValue = 99999.f;
char HISTORY_QCTEST(N_HISTORY, N_PROF, STRING16);
HISTORY_QCTEST:long_name = "Documentation of tests performed, tests failed (in hex form)";
HISTORY_QCTEST:conventions = "Write tests performed when ACTION=QCP$; tests failed when ACTION=QCF$";

HISTORY_QCTEST:_FillValue = "";

// global attributes:

:title = "Argo float vertical profile";
:institution = "";
:source = "Argo float";
:history = "2013-01-17T08:45:29Z creation";
:references = "http://www.argodatamgt.org/Documentation";
:comment = "";
:user_manual_version = "2.4";
:Conventions = "Argo-2.4 CF-1.6";
:featureType = "trajectoryProfile";
5 EASY CORA FILES NAMING CONVENTION AND FORMAT

5.1 Nomenclature of files when downloaded through the Copernicus FTP Service

INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b files nomenclature when downloaded through the Copernicus FTP service is based as follows:

-The INSITU files: ECO_code_YYYYMMDD_PR_TT.nc where:
  - code is the name of the analysis performed: DMQGCL01
  - YYYYMMDD is the date of the data,
  - PR stands for vertical Profile, TS stands for time serie
  - TT is the type of file (see 15)

Easy CORA file types content:

- **PR_PF files**: data from Argo floats directly received from the Argo DACS (real Time and delayed mode if available)\(^2\).
- **PR_BT files**: XBT and MBT measurements extracted from the CORA files
- **PR_CT files**: All the other profiles distributed from the CORA dataset
- **TS_TM files**: Timeseries from the tropical moorings (TAO/TRITON/RA3A/PIRATA moorings)
- **TS_MO files**: all the other moorings
- **TS_TS files**: TSG measurements
- **TS_DR files**: Drifters timeseries measurements

5.2 Easy CORA data and metadata

5.2.1 Best estimation of the parameters

The measured ocean parameters distributed by the Easy CORA dataset are the longitude, LONGITUDE, the latitude, LATITUDE, the time of the observation, JULD, the ocean pressure PRES or the depth, DEPTH, the in-situ temperature TEMP and the in-situ salinity PSAL. The parameters POSITION_QC, JULD_QC, TEMP_QC, PSAL_QC, PRES_QC and DEPTH_QC corresponding to the QC applied to the measurements are also distributed. In the cases were the measured parameter is the sea water pressure the corresponding depth is calculated. The sea water density is calculated with EOS80 function.

\(^2\) Principal Investigators have to provide a delayed mode validation on their floats within 6months after the date of observation
In any cases, the measured parameters only are registered in the \texttt{STATION\_PARAMETERS} metadata. The estimated \texttt{PRES} or \texttt{DEPTH} is registered in the \texttt{ESTIMATED\_STATION\_PARAMETERS} metadata.

The best estimation of the CORA parameters meets the following conditions:

- For the \texttt{TEMP}, \texttt{PSAL}, \texttt{PRES} and \texttt{DEPTH} parameters, if an adjusted version of the parameter is available (not at fillvalue), the adjusted version must be considered for the whole profile. In the other case, the original version of the parameter is considered. This method is applied independently for each parameter.
- If the temperature and the salinity parameters are both at fillvalue or with a QC at 3 or 4, the measurement must not be extracted. In the case of the measurements performed by deep ARGO (ARGO profilers, measurements below 2000m depth), if the parameter QC is at 3, if should be reset to 2 and extracted in the Easy CORA database.
- If the pressure and the depth parameters are both at fillvalue or with a QC at 3 or 4, the measurement must not be extracted.
- If the \texttt{POSITION\_QC} parameter of the \texttt{JULD\_QC} parameter are at 3 or 4, the whole profile must not be extracted.

\subsection{Profiles subsampling}

\textbf{Vertical subsampling}

A vertical subsampling is applied to the profiles in the PR files on the Easy CORA distribution. The vertical subsampling scheme follows the following recommendation.

The vertical axis is divided into bins with varying thickness from surface to bottom ocean. The thickness of the vertical bins is defined in table 7:

\begin{table}[h]
\centering
\begin{tabular}{|l|l|}
\hline
Depth/pressure & bin thickness \\
\hline
0-200 m & 1 m \\
\hline
200-1000m & 10m \\
\hline
1000- seafloor & 50m \\
\hline
\end{tabular}
\caption{Thickness of the vertical bin}
\end{table}

For each profile, a single measurement is taken in each vertical bin. The selected measurement is the first measurement contained in the bin with all parameters with a good QC.

The measurement decision criterion is defined by a weighting point-based algorithm (see table 8). Points are attributed to every measurement level contained in a bin, the measurement level with the maximum points is selected. In case of equality, the measurement level with the greater distance from the measurement of the bin located immediately below is selected.
Table 85: Scale of the points applied to estimate the relevance of a measurement level.

The points are defined in the 'Weight' column in table 8. All the points accumulated by a measurement level are summed to give the relevance of a layer.

In the case of vertical subsampling performed on a profile, the metadata `VERTICAL_SUBSAMPLING_STATUS` relative to the profile should be set to 1 (it will be set to 0 otherwise). The metadata `TEMP_SUBSAMPLING_REPRESENTATIVITY`, and `PSAL_SUBSAMPLING_REPRESENTATIVITY` must also be filled.

The `TEMP_SUBSAMPLING_REPRESENTATIVITY` is given by the distance in the mean square sense between the original profile and the subsampled profile, interpolated to the original profiles levels. This ratio varies from 0 if no subsampling has been performed to increasing positive numbers if the subsampling fails to describe the measured ocean variability.

A similar calculation is performed on the salinity measurements to fill the `PSAL_SUBSAMPLING_REPRESENTATIVITY` field.

**Time subsampling**

A time subsampling may be applied to the profiles in the PR files on the Easy CORA distribution. The aim of this time subsampling is to reduce the number of close profiles in time (and space) distributed by Easy CORA. The windowing of the time sampling as a function of the probe type is given on table 9.

<table>
<thead>
<tr>
<th>Probe type</th>
<th>Time window</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sea Glider</td>
<td>1 hour</td>
</tr>
<tr>
<td>Scanfish/Seasoar</td>
<td>1 hour</td>
</tr>
<tr>
<td>XBT</td>
<td>6 hours</td>
</tr>
<tr>
<td>CTD</td>
<td>6 hours</td>
</tr>
<tr>
<td>MBT</td>
<td>6 hours</td>
</tr>
</tbody>
</table>
### Table 9: Time window of the profile’s time subsampling as a function of the probe type.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Time Window</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sea Mammals</td>
<td>6 hours</td>
</tr>
<tr>
<td>Profilers</td>
<td>6 hours</td>
</tr>
<tr>
<td>Other instruments</td>
<td>6 hours</td>
</tr>
</tbody>
</table>

For each measuring instrument distributed in the CORA database, a single profile shall be distributed in the Easy CORA dataset, per time window defined in table 9.

The profile selection method is similar to the vertical thinning method defined in previous section. However, for the time profile selection method, the weighting points are summed over all levels in the profile and the profile with the most weight is selected. In case of equality, the profile which is most distant in time from the other profiles distributed in Easy CORA for this instrument is selected.

In the cases were a time subsampling is performed, the metadata `TIME_SUBSAMPLING_STATUS` must be set to 1 in the Easy CORA file, it will be 0 otherwise.

The `REMOVED_PROFILES_DC_REFERENCE` metadata must provide the dc_reference number associated to all the profiles removed in the time subsampling process.

### 5.2.3 Metadata distribution

The list of the metadata directly extracted from the CORA files is given below:

#### Variable: `CYCLE_NUMBER`
- **Dimensions:** N_PROF
- **Attributes:**
  - `long_name`: “Float cycle number”
  - `conventions`: “0..N, 0: launch cycle (if exists), 1, first complete cycle, etc…”
  - `FillValue`: 99999

#### Variable: `DATA_MODE`
- **Dimensions:** N_PROF
- **Attributes:**
  - `long_name`: “Delayed mode or real time data”
  - `conventions`: “R: real mode; D: delayed mode; A: real time with adjustment”
  - `FillValue`: “ ”

#### Variable: `DATA_TYPE`
- **Dimensions:** STRING16
Attributes:
  long_name: “Data type”
  FillValue: “ “

Variable: **DC_REFERENCE**
Dimensions: N_PROF, STRING32
Attributes:
  long_name: “Station unique identifier in data centre”
  FillValue: “ “

Variable: **DEPTH**
Dimensions: N_PROF, N_LEVELS
Attributes:
  long_name: “Depth under sea surface”
  standard_name: “depth”
  units: “meter”
  valid_min: 0
  valid_max: 12000
  FillValue: 99999

Variable: **DEPTH_QC**
Dimensions: N_PROF, N_LEVELS
Attributes:
  long_name: “quality flag”
  conventions: “Argo reference table 2”
  FillValue: “ “

Variable: **FLOAT_SERIAL_NUMBER**
Dimensions: N_PROF, STRING16
Attributes:
  long_name: “Serial number of the float”
  FillValue: “ “

Variable: **JULD**
Dimensions: N_PROF
Attributes:
  long_name: “Julian day (UTC) of the station relative to REFERENCE_DATE_TIME”
standard_name: “time”
units: “days since 1950-01-01 00:00:00 UTC”
conventions: “relative julian days with decimal part (as parts of day)”
FillValue: 99999
Axis: “T”

Variable: JULD_QC
Dimensions: N_PROF
Attributes:
  long_name: Quality on date and time
  conventions: “Argo reference table 2”
  FillValue: “”

Variable: LATITUDE
Dimensions: N_PROF
Attributes:
  long_name: “latitude of the station, best estimate”
  standard_name: “latitude”
  units: “degree north”
  valid_min: -90
  valid_max: 90
  axis: “Y”
  FillValue: 99999

Variable: LONGITUDE
Dimensions: N_PROF
Attributes:
  long_name: “Longitude of the station, best estimate”
  standard_name: “longitude”
  valid_min: -180
  valid_max: 180
  axis: “X”
  FillValue: 99999

Variable: PLATFORM_NUMBER
Dimensions: N_PROF, STRING8
Attributes:
long_name: “Unique platform identifier”
conventions: “WMO float identifier : A9IIIIII”
FillValue: “ ”

Variable: PLATFORM_TYPE
Dimensions: N_PROF, STRING32
Attributes:
long_name: “Type of float”
FillValue: “ ”

Variable: POSITION_QC
Dimensions: N_PROF
Attributes:
long_name: “Quality on position (latitude and longitude)”
conventions: “Argo reference table 2”
FillValue: “ ”

Variable: PRES
Dimensions: N_PROF, N_LEVELS
Attributes:
long_name: “In situ sea water pressure”
standard_name: “pres”
units: “decibars”
valid_min: 0
valid_max: 12000
FillValue: 99999

Variable: PRES_QC
Dimensions: N_PROF, N_LEVELS
Attributes:
long_name: “quality flag”
conventions: “Argo reference table 2”
FillValue: “ ”

Variable: PROJECT_NAME
Dimensions: N_PROF, STRING64
Attributes:
long_name: “Name of the project”
FillValue = “ ”
Variable: PSAL
Dimensions: N_PROF, N_LEVELS
Attributes:
  long_name: “Sea in-situ salinity”
  units: “PSU”
  valid_min: 0
  valid_max: 43
  FillValue: 99999

Variable: PSAL_QC
Dimensions: N_PROF, N_LEVELS
Attributes:
  long_name: “quality flag”
  conventions: “Argo reference table 2”
  FillValue: “ ”

Variable: STATION_PARAMETERS
Dimensions: N_PROF, N_PARAM, STRING16
Attributes:
  long_name: “List of measured parameters for the station”
  conventions: “Argo reference table 3”
  FillValue: “ ”

Variable: ESTIMATED_STATION_PARAMETERS
Dimensions: N_PROF, N_PARAM, STRING16
Attributes:
  long_name: “List of estimated parameters for the station”
  conventions: “Argo reference table 3”
  FillValue: “ ”

Variable: TEMP
Dimensions: N_PROF, N_LEVELS
Attributes:
long_name: “Sea temperature in situ”
units: “degree Celsius”
valid_min: -2
valid_max: 40
FillValue: 99999

Variable: TEMP_QC
Dimensions: N_PROF, N_LEVELS
Attributes:
  long_name: “quality flag”
  conventions: “Argo reference table 2”
  FillValue: “”

Variable: WMO_INST_TYPE
Dimensions: N_PROF, STRING
Attributes:
  long_name: “Coded instrument type”
  conventions: “Data center convention”
  FillValue: “”

Variable: PROBE_TYPE
Dimensions: N_PROF
Attributes:
  long_name: “Simplified designation of the station probe type”
  conventions: “PROBE_TYPE table in the dataset QUID”
  FillValue: 99999

  a. Easy CORA specific metadata

Variable: VERTICAL_SUBSAMPLING_STATUS
Dimensions: N_PROF
Attributes:
  long_name: “Vertical subsampling status of the station”
  conventions: “0: the profile has not been vertically subsampled; 1: the profile has been vertically subsampled”
  FillValue: “”

Variable: TIME_SUBSAMPLING_STATUS
Dimensions: N_PROF
Attributes:
  long_name: “Time subsampling status of the station”
  conventions: “0: the profile has not been subsampled in time; 1: the profile has been subsampled in time”
  FillValue: “”
Variable: **REMOVED_PROFILES_DC_REFERENCE**
Dimensions: N_PROF, STRING16, N_EV_PROF
Attributes:
  - long_name: “Removed profile DC_REFERENCE list”
  - conventions: “List the DC_REFERENCE of the profiles removed during the subsampling process. All of those profiles are in the original CORA dataset.”
  - FillValue: " "

Variable: **PSAL_SUBSAMPLING_REPRESENTATIVITY**
Dimensions: N_PROF
Attributes:
  - long_name: “Estimator of the subsampled profile salinity representativity relative to the original profile
  - conventions: “CORA QUID document”
  - FillValue: 99999

Variable: **TEMP_SUBSAMPLING_REPRESENTATIVITY**
Dimensions: N_PROF
Attributes:
  - long_name: “Estimator of the subsampled profile salinity representativity relative to the original profile
  - conventions: “CORA QUID document”
  - FillValue: 99999
6 GUIDANCE FOR THE USERS

6.1 How to use DC_REFERENCE?

Each profile has a unique identifier in the Coriolis database and the CORA dataset which is the DC_REFERENCE number. Please, refer to this DC_REFERENCE number if you want to make a feedback on a specific profile to the Coriolis data centre. The variable PLATFORM_NUMBER is the platform identifier that is assigned for the life of the platform (e.g. Profiling floats, moored buoys...). For measurements collected from research vessels or merchant ships-of-opportunity the PLATFORM_NUMBER is the vessel/ship identifier.

6.2 How to find a particular data type in CORA?

This classification of the data in netcdf files depends mainly on the data sources and resolution. However, it can be difficult for the user to find all the data from one type of instrument (e.g. CTD) as it is found in different types of files (e.g. CT, OC, TE files for CTD instruments). The variable WMO_INST_TYPE in the netcdf raw files can help to distinguish the different instrument types (see table A.1).

However the same WMO_INST_TYPE can be attributed to different types of instrument platform (e.g. the WMO_INST_TYPE 830 standing for CTD is attributed to CTD launched from vessels or ships, CTD attached to sea mammals, some mooring buoys etc...).

To facilitate the identification of a particular type of data a PROBE_TYPE code was attributed to each profile (see table A.1 for definition of codes). The PROBE_TYPE variable can be found in the Index files (see section 1.3.2) and be used to select a particular type of data.

6.3 How to find older versions of the CORA dataset?

The CORA datasets are also available on the seanoe platform, with an appropriate doi number for citation (http://www.seanoe.org/data/00351/46219/). The former version of CORA will, with time, be deleted from the Copernicus catalogue. They will thus be available on the seanoe platform.
6.4 How to select the raw or the adjusted value of a parameter?

In the CORA dataset, the parameters may be available in a regular field (TEMP for instance) and a corresponding _ADJUSTED field (TEMP_ADJUSTED in this case).

6.5 How to use Quality Flags of a parameter?

<table>
<thead>
<tr>
<th>Quality Code</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No QC was performed</td>
</tr>
<tr>
<td>1</td>
<td>Good data</td>
</tr>
<tr>
<td>2</td>
<td>Probably good data</td>
</tr>
<tr>
<td>3</td>
<td>Bad data that is potentially correctable</td>
</tr>
<tr>
<td>4</td>
<td>Bad data</td>
</tr>
<tr>
<td>5</td>
<td>Value changed (almost never used)</td>
</tr>
<tr>
<td>6</td>
<td>Not used</td>
</tr>
<tr>
<td>7</td>
<td>Not used</td>
</tr>
<tr>
<td>8</td>
<td>Interpolated value</td>
</tr>
<tr>
<td>9</td>
<td>Missing value (99999)</td>
</tr>
</tbody>
</table>

*Table 10: Quality flags and their definition*

Quality flags exist both for the PARAM (TEMP_QC, PSAL_QC,...) and the PARAM_ADJUSTED (TEMP_ADJUSTED_QC, PSAL_ADJUSTED_QC,...). Thus if one uses the adjusted values of salinity (PSAL_ADJUSTED) it should check the flag PSAL_ADJUSTED_QC to determine if the salinity value is good or not. PARAM_ADJUSTED are filled for Argo data and XBT data but for other type of platform/instrument most of the time it is an empty field.

6.6 How to use XBT with thermal offset and depth corrected?

For XBT profiles (PROBE_TYPE=10 in the index files) the corrected values are reported in the ADJUSTED fields: TEMP_ADJUSTED field for the temperature corrected from the thermal offset and DEPH_ADJUSTED for the depth corrected. The depth correction which stretches the values of depth can lead to negative depth value on the first level of some profiles (depending on the year, on the category of XBT and on the value of the first level).
Those negative depth values have been kept in the DEPH_ADJUSTED field but the quality flag has been set to 4. Those negative values concern between 50 and 70% of the profiles each year. Users are free to reject or not this first level (notice that the correction applied in the first layer is not that relevant). It appeared to us that some XBT profiles of the CORIOLIS database and CORA dataset have values of depth incorrectly stored in the field PRES. As we were not able to find the origin of this error, we computed the depth correction assuming that we had a depth and not pressure information. However to allow future corrections, for those XBT, we let the corrected depth in the PRES_ADJUSTED parameter.

6.7 How to know what is the kind of data (instrument/platform) of a given measure?

If you want a detailed (but not always filled) meta-data, you want read the variable WMO_INST_TYPE in the netcdf files that gives the instrument for each profiles.

<table>
<thead>
<tr>
<th>WMO_INST_TYPE</th>
<th>Description</th>
<th>WMO_INST_TYPE</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>Sippican T-4</td>
<td>700</td>
<td>Sippican XCTD standard</td>
</tr>
<tr>
<td>002</td>
<td>Sippican T-4 new eq.</td>
<td>710</td>
<td>Sippican XCTD deep</td>
</tr>
<tr>
<td>700</td>
<td>Sippican XCTD</td>
<td>720</td>
<td>Sippican XAABD</td>
</tr>
<tr>
<td>010</td>
<td>T-65 BR8m</td>
<td>730</td>
<td>Sippican SXCTD</td>
</tr>
<tr>
<td>021</td>
<td>Sippican Fast Deep</td>
<td>741</td>
<td>TSK XCTD</td>
</tr>
<tr>
<td>022</td>
<td>inoam002</td>
<td>742</td>
<td>TSK XCTD 2</td>
</tr>
<tr>
<td>031</td>
<td>Sippican T-6</td>
<td>743</td>
<td>TSK XCTD 2P</td>
</tr>
<tr>
<td>032</td>
<td>Sippican T-6 new eq.</td>
<td>751</td>
<td>TSK XCTD 2P</td>
</tr>
<tr>
<td>039</td>
<td>T-06-86m</td>
<td>800</td>
<td>MBT Mechanical Bath Thermograph</td>
</tr>
<tr>
<td>041</td>
<td>Sippican T-7</td>
<td>810</td>
<td>Hydrowech</td>
</tr>
<tr>
<td>042</td>
<td>Sippican T-7 new eq.</td>
<td>820</td>
<td>Thermistor Chain</td>
</tr>
<tr>
<td>049</td>
<td>T-07-86m</td>
<td>830</td>
<td>CTD</td>
</tr>
<tr>
<td>051</td>
<td>Sippican Deep Blue</td>
<td>831</td>
<td>Profiling Float (PF) - Generic</td>
</tr>
<tr>
<td>052</td>
<td>Sippican Deep Blue new eq.</td>
<td>840</td>
<td>PF, PROVOR, no conductivity sensor</td>
</tr>
<tr>
<td>059</td>
<td>T-06 760m</td>
<td>841</td>
<td>PF, PROVOR, SBR conductivity sensor</td>
</tr>
<tr>
<td>060</td>
<td>inoam000</td>
<td>842</td>
<td>PF, PROVOR, FSR conductivity sensor</td>
</tr>
<tr>
<td>061</td>
<td>Sippican T-10</td>
<td>843</td>
<td>PF, Polar Ocean Drifting System (POPS), PROVOR SBR conductivity sensor</td>
</tr>
<tr>
<td>060</td>
<td>T-10 860m</td>
<td>844</td>
<td>PF, ARVOR, Seabird conductivity sensor</td>
</tr>
<tr>
<td>071</td>
<td>Sippican T-11</td>
<td>846</td>
<td>PF, APEX, SBE conductivity sensor</td>
</tr>
<tr>
<td>079</td>
<td>T-11-860m</td>
<td>847</td>
<td>PF, APEX, FSR conductivity sensor</td>
</tr>
<tr>
<td>081</td>
<td>Sippican AXBT (800m probes)</td>
<td>850</td>
<td>PF, SOLO, no conductivity sensor</td>
</tr>
<tr>
<td>201</td>
<td>TSK T-4</td>
<td>851</td>
<td>PF, SOLO, SBE conductivity sensor</td>
</tr>
<tr>
<td>202</td>
<td>TSK T-4 new eq.</td>
<td>852</td>
<td>PF, SOLO, FSR conductivity sensor</td>
</tr>
<tr>
<td>211</td>
<td>TSK T-6</td>
<td>853</td>
<td>PF, SOLO/SCRRPSS, Seabird conductivity sensor</td>
</tr>
<tr>
<td>212</td>
<td>TSK T-6 new eq.</td>
<td>855</td>
<td>PF, NINJA, SBR conductivity sensor</td>
</tr>
<tr>
<td>221</td>
<td>TSK T-7</td>
<td>856</td>
<td>PF, NINJA, TSK conductivity sensor</td>
</tr>
<tr>
<td>222</td>
<td>TSK T-7 new eq.</td>
<td>860</td>
<td>PF, DEMO, SBE conductivity sensor</td>
</tr>
<tr>
<td>254</td>
<td>TSK T-5</td>
<td>861</td>
<td>PF, DEMO, FSR conductivity sensor</td>
</tr>
<tr>
<td>251</td>
<td>TSK T-5</td>
<td>995</td>
<td>Instrument attached to marine mammals</td>
</tr>
<tr>
<td>259</td>
<td>TSK T-10</td>
<td>999</td>
<td>Unknown</td>
</tr>
<tr>
<td>253</td>
<td>TSK Deep Blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>272</td>
<td>TSK Deep Blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>401</td>
<td>Spartan XBT-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>411</td>
<td>Spartan XBT-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>412</td>
<td>Spartan XBT-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>451</td>
<td>Spartan XBT-6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>460</td>
<td>Spartan XBT-7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>462</td>
<td>Spartan XBT-7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>481</td>
<td>Spartan XBT-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>491</td>
<td>Spartan XBT-10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>501</td>
<td>Spartan XBT-30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Otherwise you can have more general information using the PROBE_TYPE variable from Index files:
<table>
<thead>
<tr>
<th>PROBE_TYPE</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>XBT</td>
</tr>
<tr>
<td>20</td>
<td>CTD</td>
</tr>
<tr>
<td>30</td>
<td>XCTD</td>
</tr>
<tr>
<td>40</td>
<td>PROFILING FLOAT</td>
</tr>
<tr>
<td>51</td>
<td>TAO-TRITON PIHATA RAMA MOORINGS</td>
</tr>
<tr>
<td>52</td>
<td>COASTAL MOORINGS (&lt; 20km from the coast)</td>
</tr>
<tr>
<td>50</td>
<td>OTHER MOORINGS</td>
</tr>
<tr>
<td>60</td>
<td>GLIDERS</td>
</tr>
<tr>
<td>70</td>
<td>INSTRUMENT ATTACHED TO SEA MAMMALS</td>
</tr>
<tr>
<td>80</td>
<td>DRIFTING BUOYS</td>
</tr>
<tr>
<td>0</td>
<td>UNKNOWN</td>
</tr>
</tbody>
</table>