The reference dataset in the Mediterranean and Black Seas for DMQC activity

by

G. Notarstefano

Deliverable 4.4.1 of the EU MOCCA project

Approved for release by: ……

Dr. Cosimo Solidoro
Director, Department of Oceanography

Table of Contents

Table of Figures

1. INTRODUCTION

This document describes the activities related to the conductivity-temperature-depth (CTD) reference dataset used for the Delayed-Mode Quality Control (DMQC) of the MOCCA fleet in the Mediterranean and Black Sea.

The DMQC of MOCCA floats is performed by Euro-Argo MOCCA partners delayed-mode operators according to the area of deployment and taking into account their area of expertise. It is further described in the deliverable D4.1.1 Organization of Float Data Management among DAC and DM-operators.

OGS is responsible for the DMQC of the temperature and salinity data of the MOCCA floats deployed in these two marginal seas.

The DMQC processing is described in the deliverable D.4.3.1 Report on Delayed-Mode processing on the MOCCA fleet.

The following map (Figure 1) illustrates the MOCCA floats locations in the Mediterranean and Black Sea:

Figure 1: MOCCA floats locations in the Mediterranean and Black Seas. Dots represent the last platform location as of 18th January 2019 (active floats in green, inactive floats in red).

2. REFERENCE DATASET

Argo floats are usually not recovered for calibration purposes and hence the delayed-mode technics for checking the conductivity sensor drifts and offsets are based on reliable CTD reference datasets.

Two different CTD reference datasets are provided and updated on a regular basis by designated figures. One reference dataset consists of CTD data collected worldwide by several research institutes (hereby referred as to **CTD reference dataset**). Data are selected, aggregated and quality controlled as explained in the Argo quality control manual for CTD and trajectory data, version 3.1 (*http://dx.doi.org/10.13155/33951).* Once a new version of the CTD reference dataset is ready, it is available for downloading at a dedicated FTP address for all the DMQC operators. The second reference dataset is built using the Argo CTD profiles that respect a set of predefined criteria (hereby referred as to **Argo CTD reference dataset**), as stated in the Argo quality control manual for CTD and trajectory data, version 3.1.

This report focuses on the use of these two datasets in a marginal sea context like the Mediterranean and Black Seas. Moreover, it describes the work done at regional level to try to improve the official CTD reference dataset with other new and more updated CTD data.

2.1. Regional needs

The Mediterranean Sea is characterized by a complex bathymetry, where the existing shallow areas represent a threshold in the selection criteria for the CTD profiles to be retained in the official reference database (only profiles that sampled deeper than 900 dbar are selected). Moreover, some local research institutes that collect CTD data at regular basis don't share they data because they are not part of dedicated infrastructures or international projects. This can cause the CTD reference dataset not to be updated and scarce in temporal and spatial coverage in some areas. Since the Mediterranean Sea is characterized by several water masses that can change properties dramatically over the years, it is crucial to have the best co-location (in space and time) between the CTD reference dataset and the Argo CTD profiles in order to separate differences between the two datasets due to sensor drift or to the change of water mass properties. For these reasons, OGS, as responsible of the DMQC activities in the Mediterranean and Black Sea, tries to collect CTD data in complement of the official CTD reference dataset using mainly two approaches: personal contacts from one side and regional data services from another side.

2.2. CTD reference dataset

The CTD reference dataset is provided by the Coriolis in-situ Service for operational oceanography and it is usually updated once or twice a year. At the time of this report, the latest version available is "CTD_for_DMQC_2018V01". Files are organized following the World Meteorological Organization (WMO) boxes denomination that consists of geocodes that divides the chart of the world into 10° latitude by 10° longitude grid cells (Figure 2).

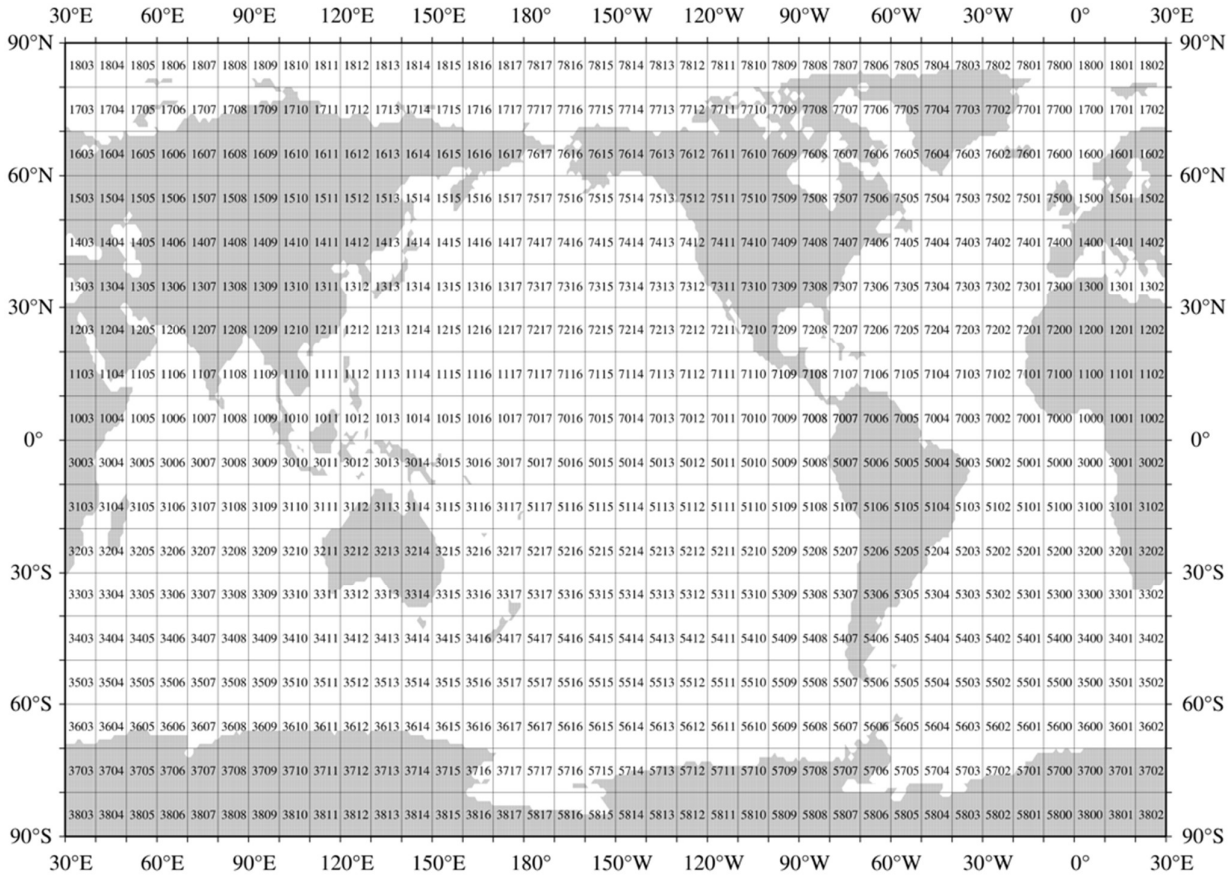

Figure 2: World Meteorological Organization (WMO) 10-degree squares.

The 10-degree squares that cover the Mediterranean and Black Sea area are the following:

1300, 1301, 1302, 1303, 1400, 1401, 1402, 1403, 1404, 7300, 7400

Hence, the above files are downloaded by the Coriolis FTP address, uncompressed and saved in MatLab (.mat) format. The temporal and spatial distribution of these CTD profiles is shown in Figure 3 and Figure 4, respectively.

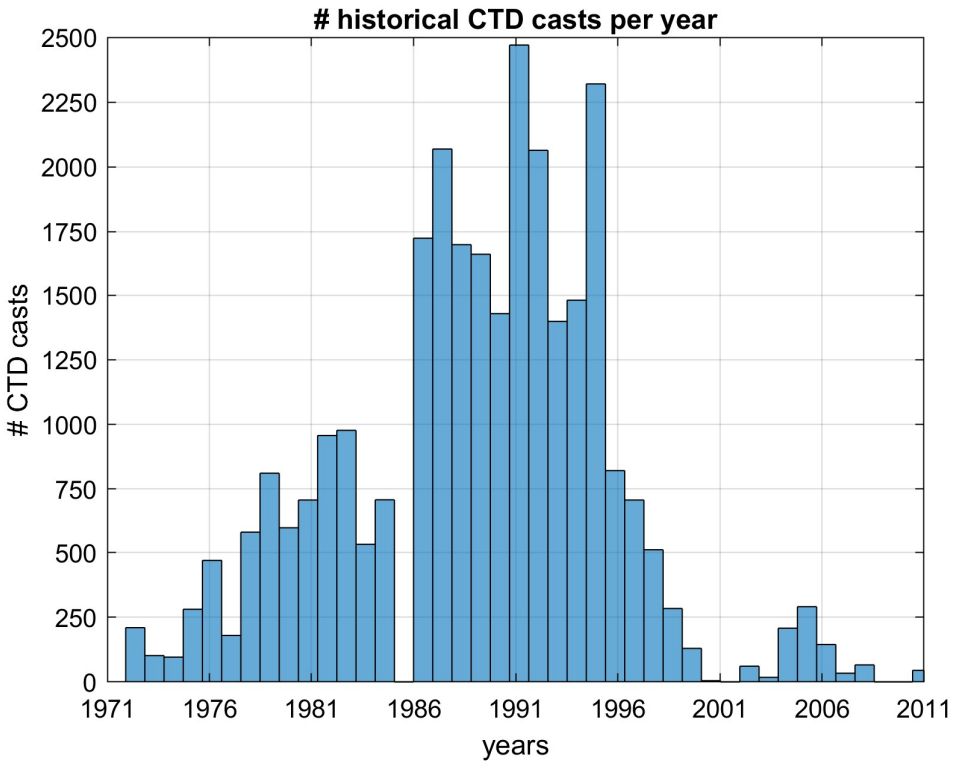

Figure 3: temporal distribution of the CTD profiles available in the CTD reference dataset version "CTD_for_DMQC_2018V01".

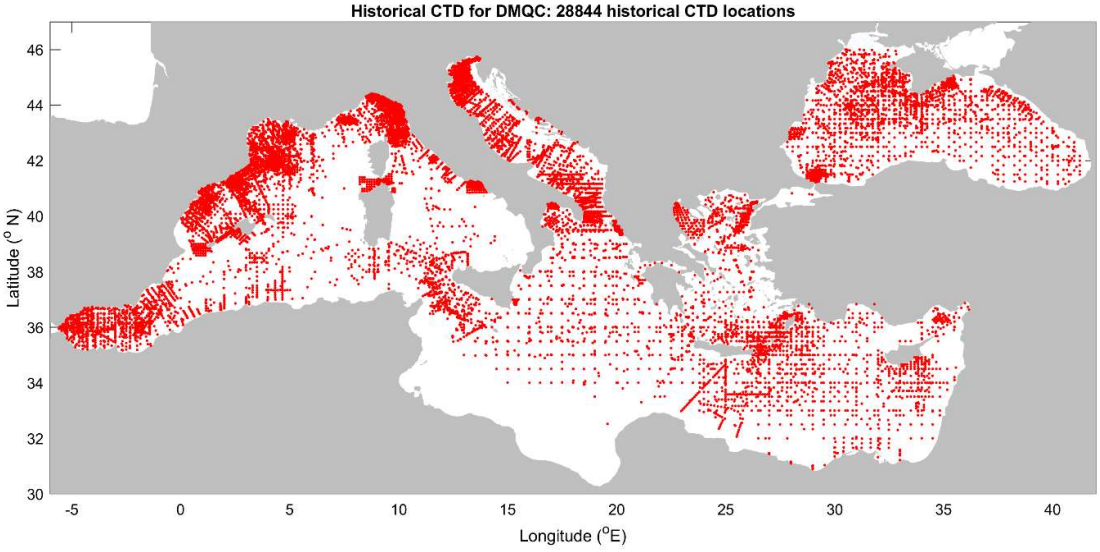

Figure 4: spatial distribution of the CTD profiles available in the CTD reference dataset version "CTD_for_DMQC_2018V01".

The temporal distribution exhibits a consistent increase from about 1972 till the year 1995 and then a sharp decrease is evident; the last CTD profiles available are in 2011. Some areas are highly under sampled, like the Southern Ionian and the Thyrrhenian Sea.

2.3. Argo CTD reference dataset

The Argo CTD reference dataset is also provided by the Coriolis in-situ service for operational oceanography and files are organized in the same way as for the CTD reference dataset. The latest version available at Coriolis is "ARGO_for-DMQC_208V01".

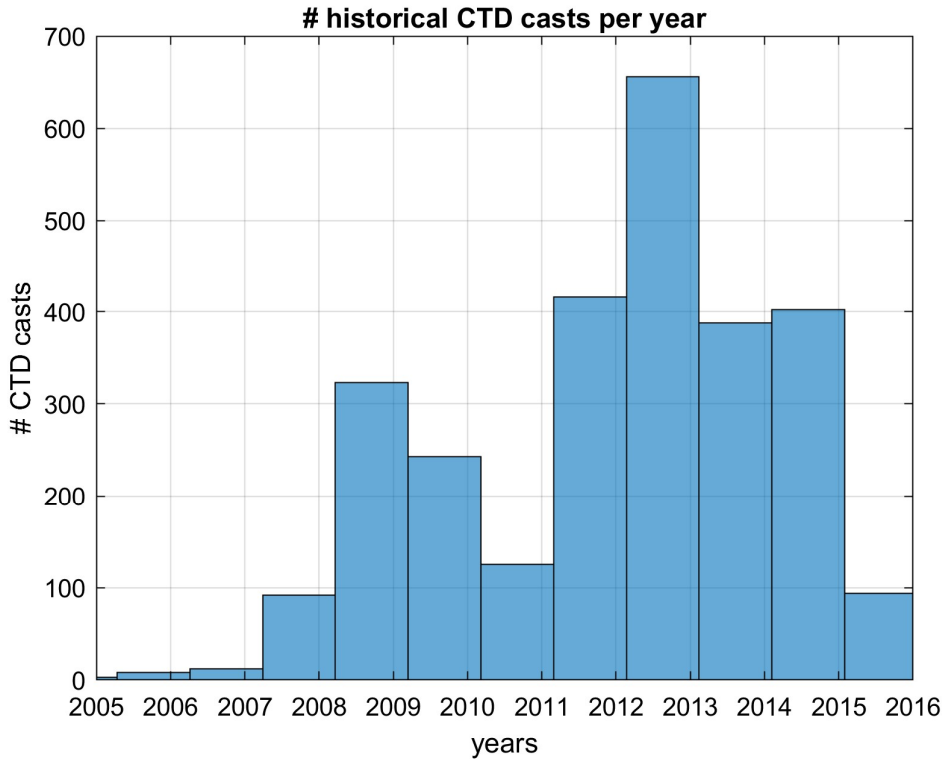

Figure 5: temporal distribution of the Argo CTD profiles available in the Argo CTD reference dataset version "ARGO_for_DMQC_2018V01".

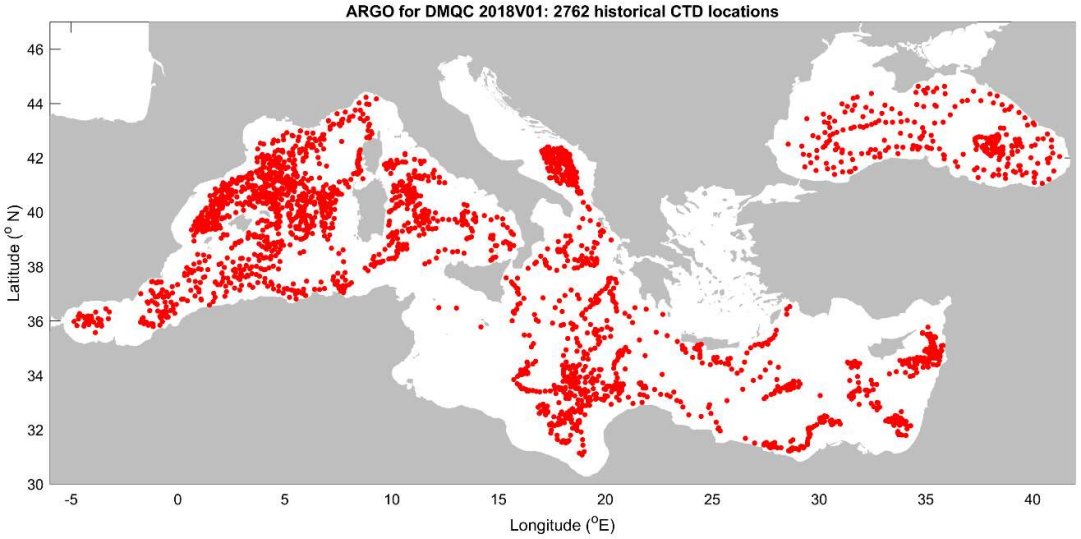

Figure 6: spatial distribution of the Argo CTD profiles available in the Argo CTD reference dataset version "ARGO_for_DMQC_2018V01".

The Argo CTD profiles are quite well homogeneously distributed, both in time and space, in the several sub-basins of the Mediterranean Sea and also in the Black Sea (Figure 5 and 6). Shallow areas, where Argo floats are not often deployed, don't have an accurate coverage (Sicily Channel, Aegean and Middle-North Adriatic).

3. IMPROVING OF THE CTD REFERENCE DATASET

Since the DMQC activity requires the availability of a good reference dataset, a work was conducted to explore the possibility to fill some gaps (in time and space) in the CTD reference dataset.

3.1. CTD profiles obtained through personal contact

The collection of CTD profiles through personal contacts started in 2008 and since then several profiles have been used to improve the CTD reference dataset. European colleagues from different research institutes kindly provide us with CTD data acquired during regular cruises or in framework of projects. A lot of work has been done in finding the right contacts, email exchanging and gathering the data. Dozens of datasets were collected in this way spanning from 1997 to 2017 and from the Alboran to the Levantine Seas. The CTD profiles used to build the MEDAR-MEDATLAS climatology have also been added to these datasets and it consists of data from 1972 to 2000.

The last CTD data collection was done in the second part of 2018 under the MOCCA project activity and it consists of CTD profiles from 2013 to 2017 in the Adriatic, Alboran, Algerian, Ionian, Thyrrhenian Seas, Sicily Channel and Cretan Passage.

The files were usually received in different formats and hence file-reading MatLab scripts were prepared accordingly. The data were supposed to be already of good quality but a light additional quality control has been applied in order to remove any residual outliers and spikes.

Many CTD data collected were policy free and hence they have been shared with the Coriolis in-situ Service and integrated in the “official” CTD reference dataset. The data policy is discussed with the owner of the data and as soon as they are declared “not restricted”, they become part of the “official” dataset.

The CTD profile locations of this dataset is shown in Figure 7 and the respective temporal distribution is in Figure 8.

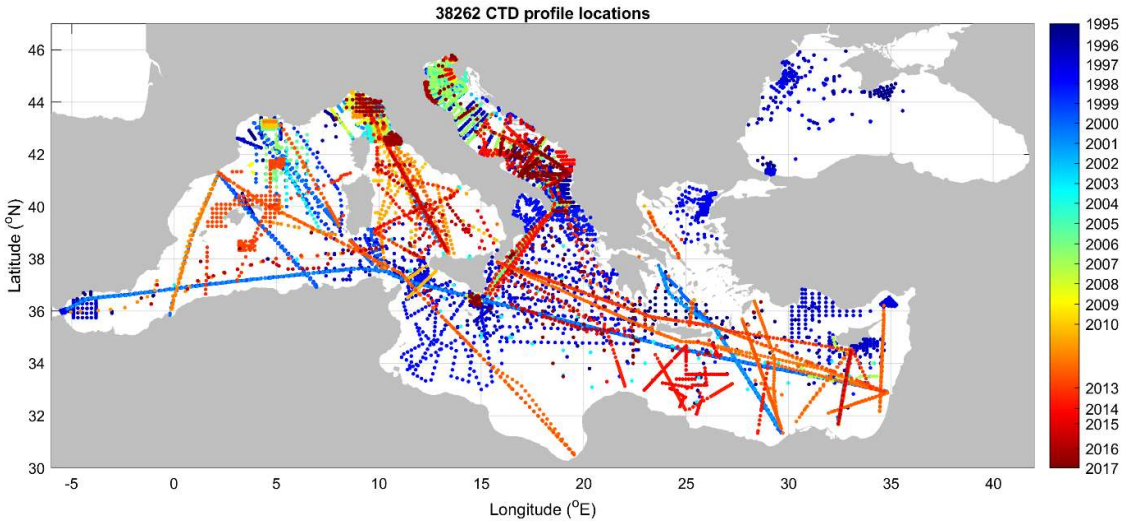

Figure 7: spatial distribution, color-coded for time, of the CTD profiles collected and used as a complement of the CTD reference dataset.

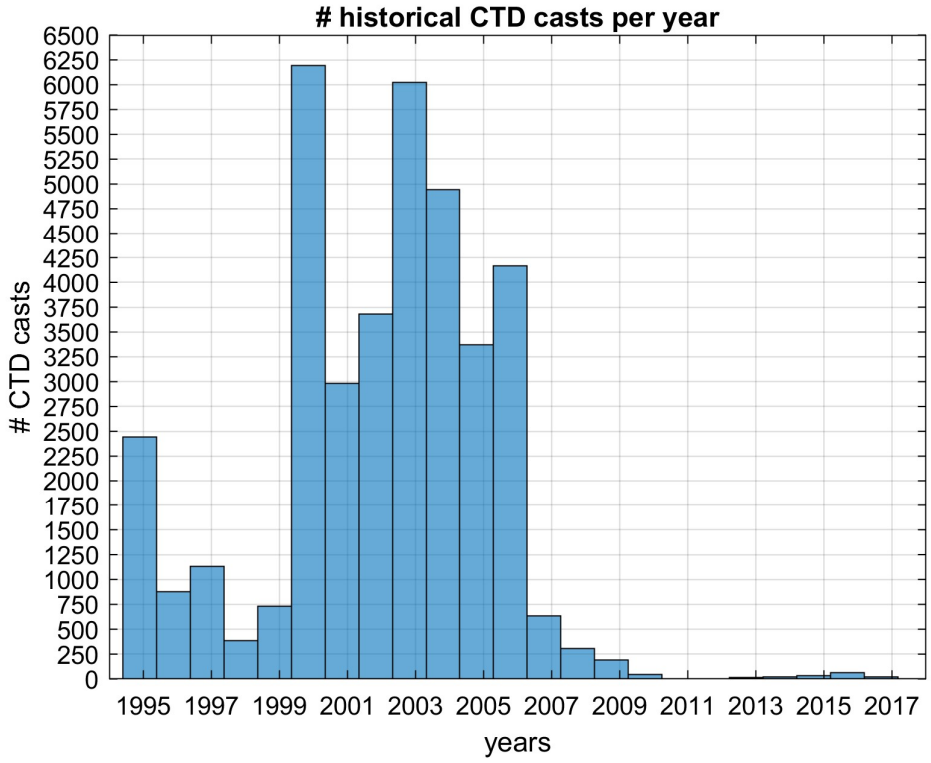

Figure 8: temporal distribution of the CTD profiles collected and used as a complement of the CTD reference dataset.

3.2. CTD profiles obtained through dedicated services

Another CTD data source taken into consideration is the one connected to the marine monitoring services. The Copernicus Marine Environment Monitoring System (CMEMS) has been chosen because it provides a great quantity of data that follow a multiple level quality control procedure. These data are not integrated in the CTD reference dataset by the Coriolis in-situ Service. The CMEMS files are in NetCDF format and are available through a dedicated FTP server. The CTD files have to be extracted by a common folder reading the file name that is coded per platform and profile types. MatLab scripts have been built to convert the files from NetCDF to MatLab format. The procedure has been done twice in two different data repository: the first time to collect the CTD data in the Mediterranean Sea (Figure 9 and 10) and the second time for the Black Sea (Figure 11 and 12).

The last CTD data collection has been done in 2018 in the framework of the MOCCA project and the following two CMEMS products were used:

- INSITU_MED_TS_REP_OBSERVATIONS_013_041
- INSITU_BS_TS_REP_OBSERVATIONS_013_042

for the Mediterranean and Black Sea, respectively.

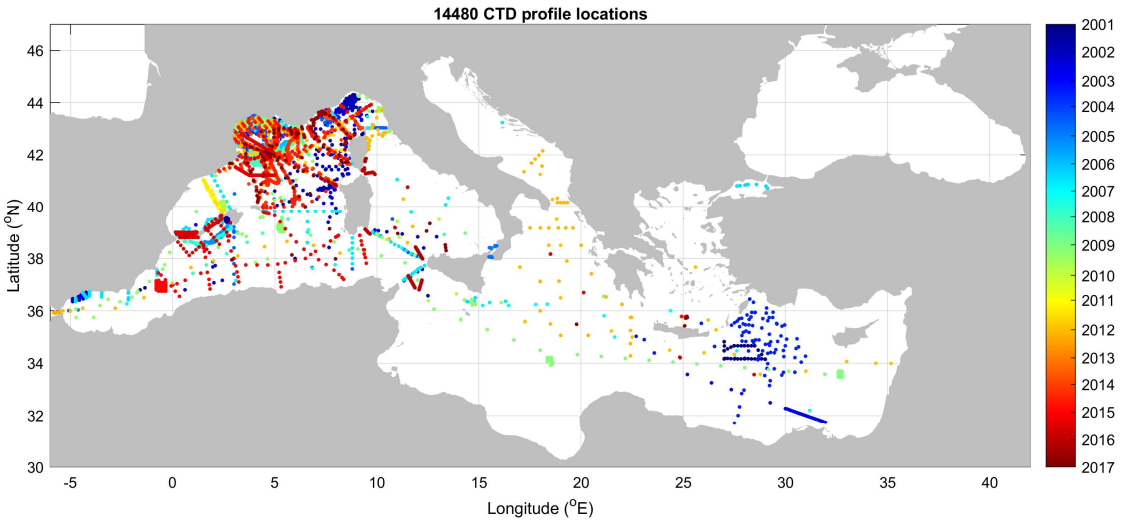

Figure 9: spatial distribution, color-coded for time, of the CTD profiles collected through the CMEMS portal and used as a complement of the CTD reference dataset of the Mediterranean Sea.

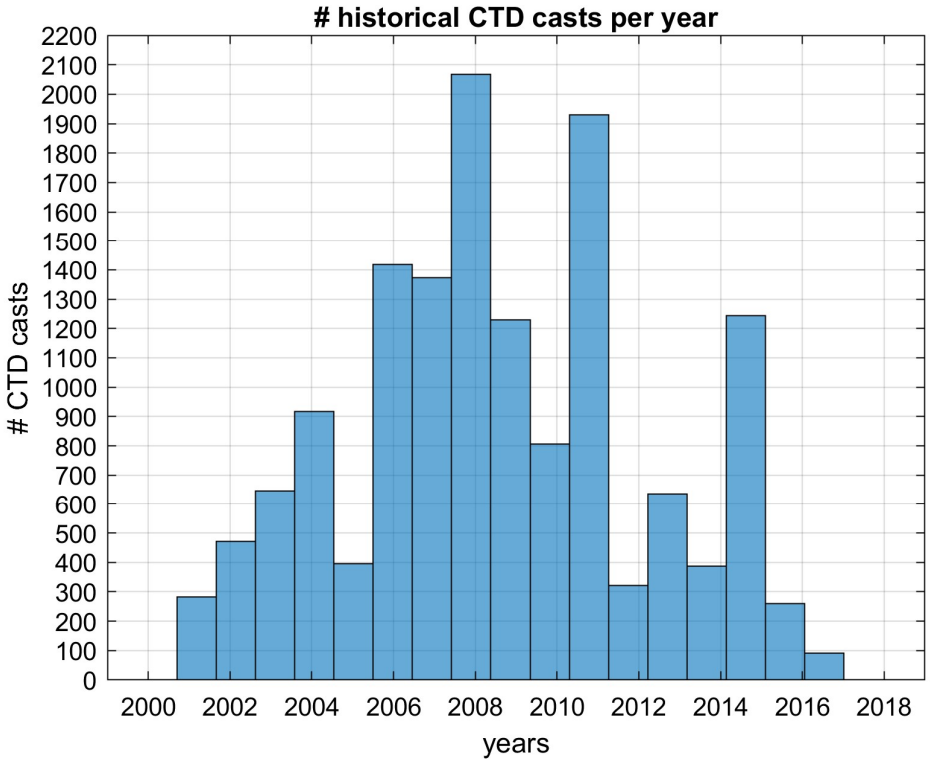

Figure 10: temporal distribution of the CTD profiles collected through the CMEMS portal and used as a complement of the CTD reference dataset of the Mediterranean Sea.

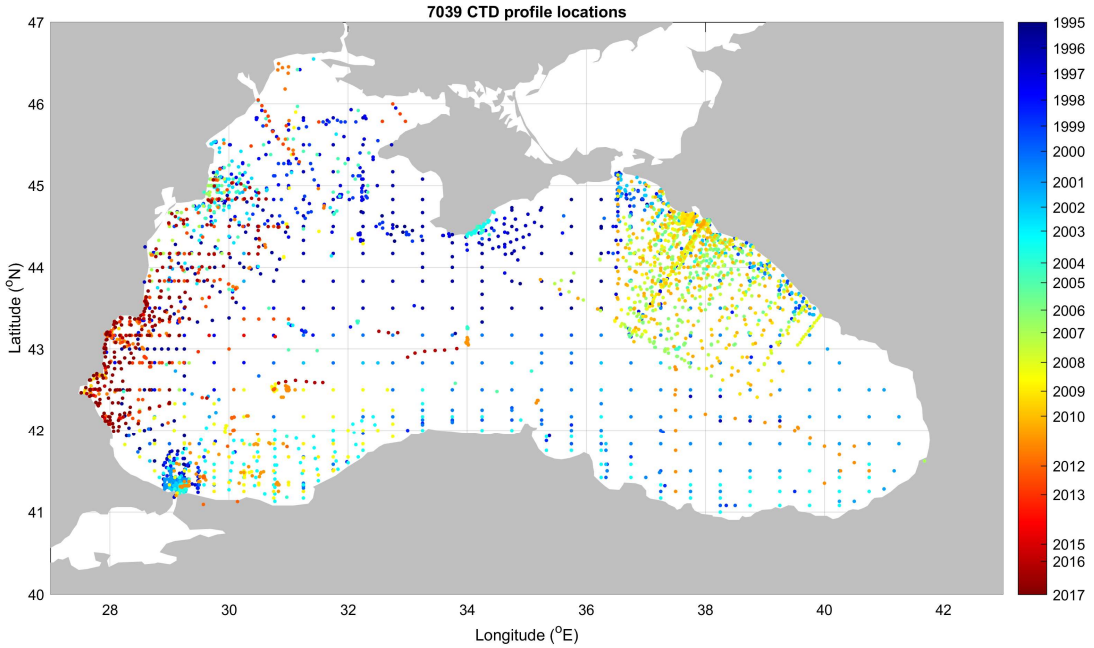

Figure 11: spatial distribution, color-coded for time, of the CTD profiles collected through the CMEMS portal and used as a complement of the CTD reference dataset of the Black Sea.

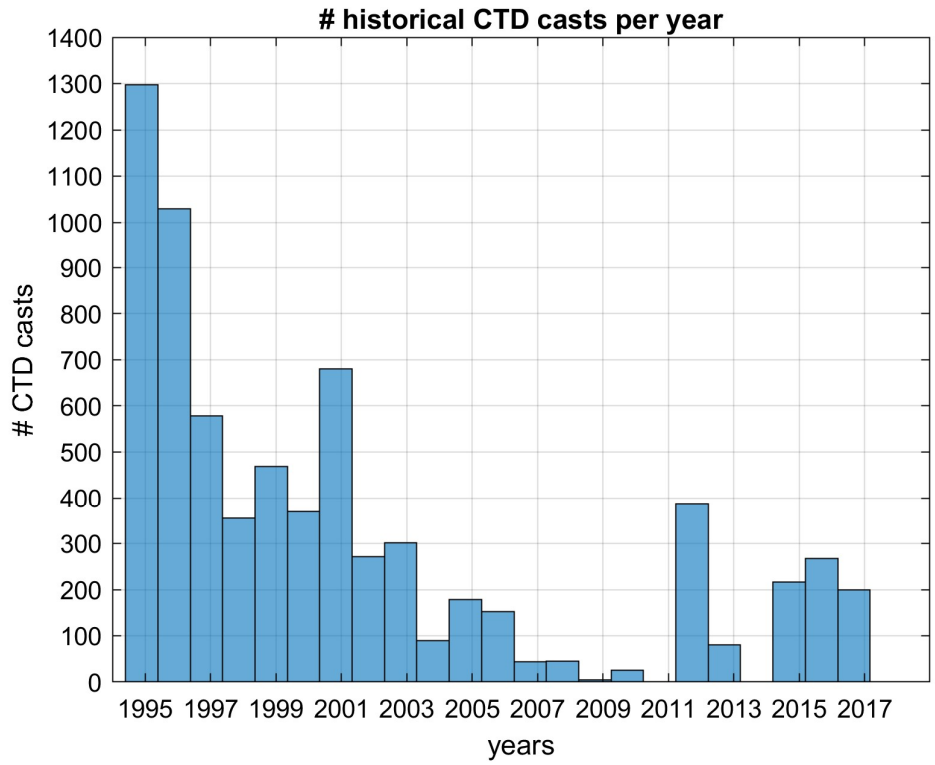

Figure 12: temporal distribution of the CTD profiles collected through the CMEMS portal and used as a complement of the CTD reference dataset of the Black Sea.

4. MERGER OF THE DATASETS

4.1. Checking for duplicates

Once the files are converted from their original format into MatLab format, a further step is requested to make them compatible to be used by the Owens-Wong (OW) software for calibration purposes. The CTD data are then compared to the CTD reference dataset and checked to remove duplicates taking into account thresholds of 10 minutes and 100 meters for time and location respectively. CTD profiles whose difference in time and space is less than the above predefined thresholds are considered duplicates and hence removed from the dataset.

4.2. Subset of the CTD data and merger

The CTD data are first separated into 10° X 10° WMO boxes. Then, due to the different nature of the existing water masses and to the geography of the Mediterranean Sea, the CTD data within the WMO boxes are grouped according to the dimension of various climatological sub-basins, as defined by the EU/MEDAR-MEDATLA II project (Figure 13).

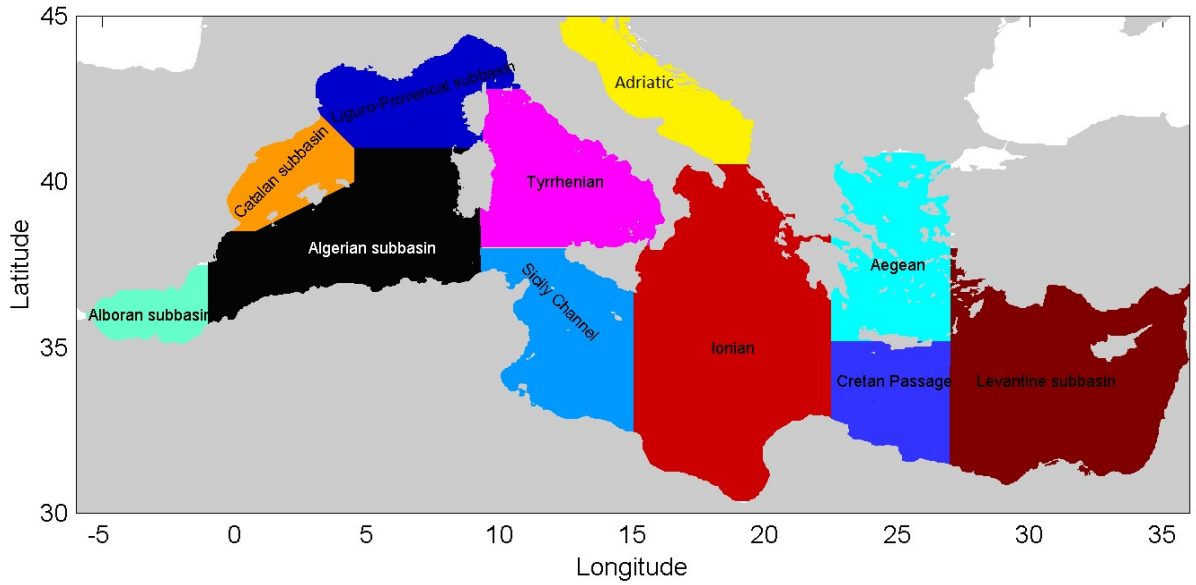

Figure 13: sub-basins of the Mediterranean Sea, based on the climatological areas defined by the EU/MEDAR-MEDATLAS II project.

The new CTD data are eventually merged into the CTD reference dataset and this final version is summarized in Figure 14 and 15. The dataset consists of about 56000 CTD profiles. Data before year 1995 were discarded because considered too old for quality control purposes.

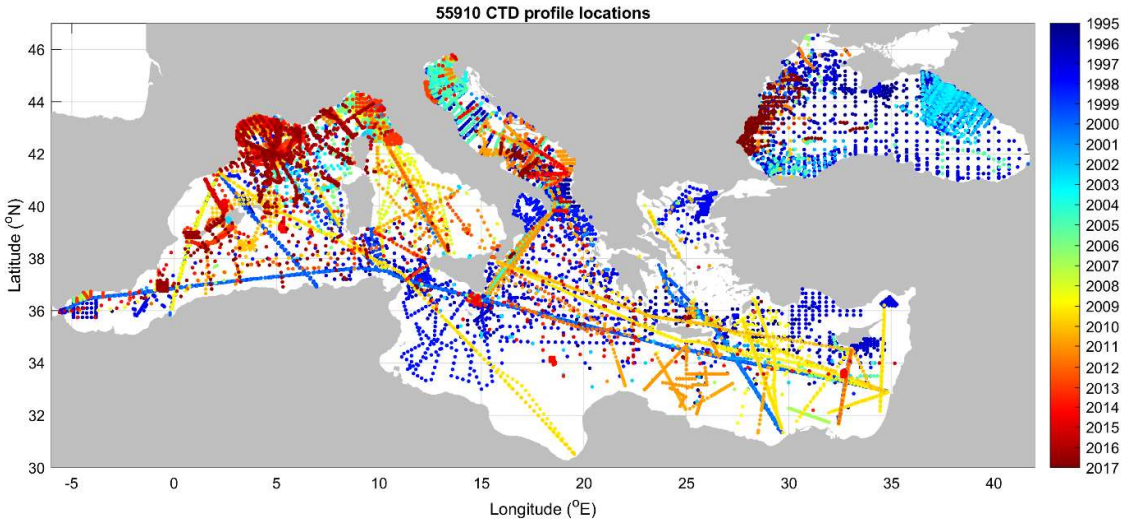

Figure 14: spatial distribution, color-coded for time, of the CTD profiles in the final version of the CTD reference dataset of the Mediterranean and Black Seas.

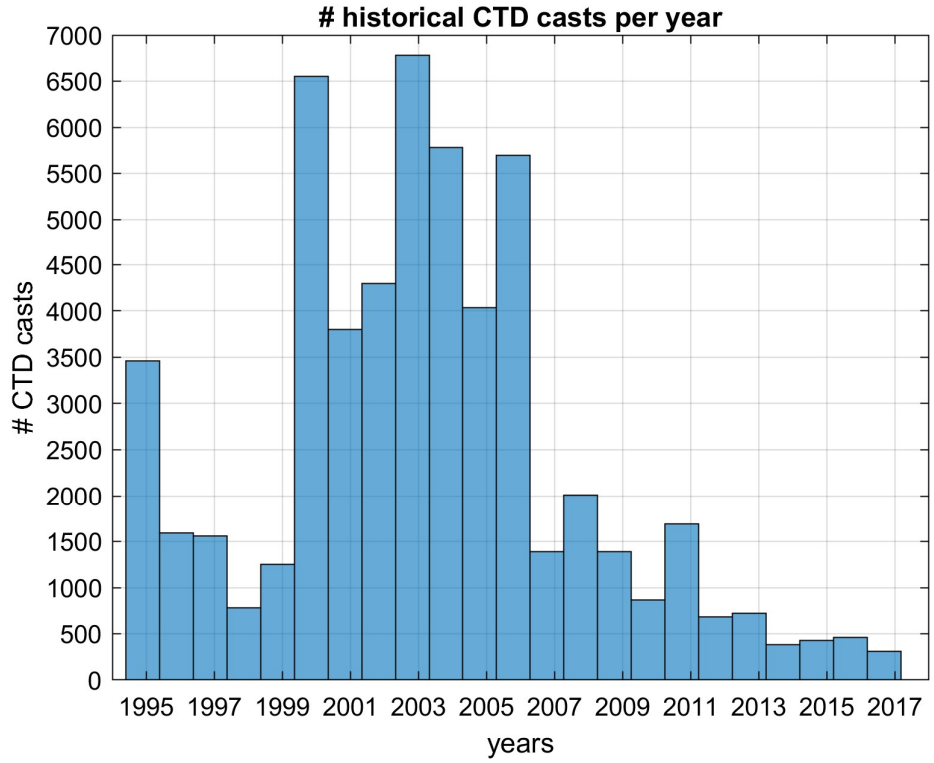

Figure 15: temporal distribution of the CTD profiles in the final version of the CTD reference dataset of the Mediterranean and Black Seas.

5. CONCLUSION

A great effort has been required to gather additional CTD data from different sources as a complement of dataset provided by the Coriolis in-situ Service. A consistent improvement of the CTD reference dataset has been done for the Argo DMQC activity in the framework of the MOCCA project. The policy of the aggregated data will be further discussed in order to make the data available through the Coriolis in-situ Service for DMQC activities only.