
TAIPro2022: the Tyrrhenian Sea & Algero-Provencal component of the MedSHIP programme

A Data Management Plan created using DMP Roadmap for Eurofleets+

Creator: Katrin Schroeder

Affiliation: Your Research Institute

Template: Eurofleets+

ORCID ID: <https://orcid.org/0000-0001-7991-9121>

Project abstract:

The Mediterranean ocean observing community is committed to conduct regular repeat hydrological surveys through the MedSHIP program, thereby sampling the entire Mediterranean Sea in a systematic way (see Schroeder et al., 2015, Oceanography). Here we propose to repeat the two western Mediterranean meridional sections in an 11-day expedition on board RV BELGICA II during summer of 2022 (cruise TAIPro2022). The last occupation of these transects was conducted in 2016 during the EF cruise TAIPro2016 (Jullion, 2016), to which most participants to this proposal contributed. The primary purpose of these regular repeats is to quantify and characterize changes in the Mediterranean sea on a sub-decadal time frame. We propose to do so by conducting high-accuracy observations of all “level 1” variables defined in the GO-SHIP program (www.goship.org) and compare those to historical data. In addition we will quantify transport and properties of the flow through the straits between western Mediterranean, Tyrrhenian Sea and Eastern Mediterranean Sea. This program complements the ship-based component of the MOOSE program in that the geographical range is larger and we plan to measure additional variables normally not included in the MOOSE program.

Last modified: 18-03-2021

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

TAIPro2022: the Tyrrhenian Sea & Algero-Provencal component of the MedSHIP programme - Phase 1: Preliminary DMP

1. Data Summary

The data collected will be physical thermohaline properties of water masses, biogeochemical properties (inorganic nutrients, oxygen, inorganic carbon, transient tracers, radionuclides). These that will allow for, when compared to previously collected data, to meet to MedSHIP objectives, i.e. to assess changes: heat and fluxes changes, circulation and mixing, transient tracers, anthropogenic carbon and acidification, oxygen and deoxygenation, nutrients, multi-variable analysis. Additionally, time series of temperature, salinity, dissolved oxygen and currents will be downloaded from the instruments of three moorings, in order to assess the role of inflow and outflow of the intermediate water in determining observed changes in the Tyrrhenian Sea.

Zooplankton species in relation to environmental variables provide a global vision of the evolution of their biodiversity, and ultimately of the consequences on the biogeochemical and trophic functioning of the pelagic ecosystem. Analysis from imagery system like UVP5 (Under Vision Profiler) offers a map of the distributions of biomass and global zooplankton abundance as well as taxonomic groups.

Data format for CTD will be standard *.cnv, ODV and *.mat files, LADCP data will be generated as *.mat files for internal use, and as ASCII files for distribution, chemical data (bottle data) will also be organized in WOCE exchange format. Mooring data will be collected in netcdf format. Zooplankton and particles abundance data are stored in .txt format and can be easily imported to ODV software. UVP5 pictures are identified automatically and, after being exported in different formats, stored to ECOTAXA database (<https://ecotaxa.obs-vlfr.fr>).

Automatic systems, and near-real time post-processing on board, will give the possible to have ready-to-deliver data, that will be sent to shore (Coriolis system, or analogous) via email. This might include CTDs and underway data. Data that require more time-consuming post-processing and data coming from chemical analysis will not be transferred to shore during the cruise. Particles abundance from UVP5 will be processed onboard and zooplankton identification will be done in the lab.

DATA TYPE/PROCESSING	QA	DATA HARMONISATION
CTD	outlier handling, missing and suspect values	consistent use of headings and data formatting
BIOGEOCH. DATA	outlier handling, missing and suspect values	QC flagging
CHEMIC. DATA	outlier handling, missing and suspect values	QC flagging
PARTICLES DATA	Process size spectra from the particle data (outlier handling, missing and suspect values)	QC flagging
ZOOPLANK. DATA	QA with automatic recognition process	identified by taxonomist experts

These steps will be performed after the cruise and (for bottle data) after all the chemical analysis has been concluded. May take up to 6 months for physical data and up to 1 year for bottle data.

2.1 Making data findable, including provisions for metadata

Mediterranean Sea, hydrographic transects, MedSHIP, GO-SHIP, physical data, biogeochemical data, transient tracers

All information will be added to the cruise report, in specific subchapters dedicated to the single data types. This cruise report will be continuously updated, sharing an online editable document among all partners, until the last data are ready and postprocessed. It will then be the source of information to create the correct metadata elements when submitting the

data to the Emodnet Ingestion Portal.

2.2. Making data openly accessible

All datasets will be shared according to the timeline defined by GOSHIP procedures (mentioned also in the project proposal). According to the recommendations given by CIESM, the data-release guidelines of the GOSHIP Program (IOC, 2009) will be adopted, in order to be compliant with the global program requirements. While it is important to protect individual scientific interest and investment of effort by investigators, it is necessary to evolve towards a more operational system, to justify a sustained program with national or international funding support. A closer cooperation with other operational programs must require specific data-release practices. Thus, TAIPro-2021 will adopt the following:

- Preliminary data set released within 6 weeks (e.g. all data measured on the ship)
- 6 months for final physical data
- 1 year for final data of all other variables (except for isotopes or tracers)

The Final Cruise Results will be promptly released in form of reports on the data (in journals like, for example, Earth System Science Data, ESSD) and through publications in scientific journals, possibly in conjunction with the other Med-SHIP initiatives carried out.

We don't have any experience with EuroFleets data repository, but would be happy to submit all data to SeaDataNet or EMODNET for repository. All data from this cruise will be published according to the GO-SHIP recommendations, which are listed at <http://www.goship.org/DataDirect.html>. In particular, CTD and bottle data will be sent to CCHDO (CLIVAR and Carbon Hydrographic Data Office), and will be stored at NCEI in addition.

2.3. Making data interoperable

Cruise dataset will be stored in SISMER database and metadata and dataset from CTD profiles and bottle sampling will be stored in SEANOE providing a list of authors and DOI for future publications. The dataset will be formatted in NetCDF which the most interoperable format. The data management units where we will submit our data will make data interoperable, that is the responsibility of the data portals.

2.4. Increase data re-use (through clarifying licences)

CC BY 4.0

TAIPro2021 will adopt the following:

- Preliminary data set released within 6 weeks (e.g. all data measured on the ship)
- 6 months for final physical data
- 1 year for final data of all other variables (except for isotopes or tracers)

All data will be free and open access without restrictions as outlined in the GO-SHIP data management guidelines.

4. Data security

Multiple backups, including storage on institutional Clouds, are foreseen for all data that will be shared among the partners of the cruise.

In particular, the Kiel data management team of GEOMAR will support this cruise already after the assignment of ship time. This means that all general cruise information is freely available from the Kiel Ocean Science Information System (OSIS). This basic information will be extended right after the cruise is finished with the scientific sampling events and general cruise underway data.

The raw data acquired during the cruise (see section 3.2) are publicly visible and stored as georeferenced data sets within OSIS. This system stores metadata supplied by the researchers directly along with scientific data. The predefined process of cruise program publication with the ability to reference data files by sampling events enables merging of scientific raw

and primary data collections. This first level data will be synchronized subsequently with the GEOMAR and data base system allowing for an immediate exchange of the data within the project group. Subsequent results may be added in the same context, as well as print publications related to the cruise or scientific measurements derived from this cruise. The GEOMAR library connects the publication repository with OSIS for overarching cruise output reporting. The direct publication of the metadata promotes communication with scientist outside the project and institutes without endangering the safety of the scientific data.

The metadata will be made publicly accessible immediately. Data exchange before this date between us is desired and will take place as required. Published data will be submitted to World Data Centers (WDC). The GEOMAR data management team will take care of the data transfers to long-term archives in order to ensure the data availability worldwide and for the far future.

6. Other issues

See above, and the GO-SHIP data management guidelines.