Data Science Journal, Volume 13, 27 January 2015

IBAMAR DATABASE: FOUR DECADES OF SAMPLING ON THE WESTERN MEDITERRANEAN SEA

***A Aparicio-González*[1*]*, J L López-Jurado*[1]*, R Balbín*[1]*, J C Alonso*[3]*, B Amengual*[1]*, J Jansá*[1]*, M C García*[2]*, F Moyá*[2]*, R Santiago*[1]*, M Serra*[1]*, M Vargas-Yáñez*[2]**

[1]*Instituto Español de Oceanografía, Centro Oceanográfico de Baleares, Palma de Mallorca, Baleares, Spain*
**Email:*alberto.aparicio@ba.ieo.es
[2]*Instituto Español de Oceanografía, Centro Oceanográfico de Málaga, Málaga, Spain*
[3] *Consejo Superior de Investigaciones Científicas, Instituto Mediterráneo de Estudios Avanzados, Esporles, Baleares, Spain*

ABSTRACT

IBAMar is a regional database that puts together all the physical and biochemical data provided by multiparametric probes and water sample analysis taken during the cruises managed by the Balearic Oceanographic Center of the Instituto Español de Oceanografía (COB-IEO) during the last four decades. Initially, it integrated data from hydrographic profiles obtained from CTDs (conductivity, temperature, depth) equipped with several sensors, but it has been recently extended to incorporate data obtained with hydrocasts using oceanographic Niskin or Nansen bottles. The result is an extensive regional resource database that includies physical hydrographic data such as temperature (T), salinity (S), dissolved oxygen (DO), fluorescence, and turbidity, as well as biochemical data, specifically dissolved inorganic nutrients (phosphate, nitrate, nitrite, and silicate) and chlorophyll-a. Different technologies and methodologies were used by independent teams during the four decades of data sampling. However in the IBAMar database, data have been reprocessed using the same protocols and a standard quality control (QC) methodology has been applied to each variable. The result is a homogeneous and quality-controlled data. IBAMar database at standard levels is freely available for exploration and download from http://www.ba.ieo.es/ibamar/.

1 INTRODUCTION

Oceanographic databases are tools for preserving historical and current observations under a common structure facilitating the characterization of a marine environment, including its variability and the geographical distribution of the parameters. Whenever the spatio-temporal distribution of observations are sufficient, databases are also useful in detecting oscillations and trends in oceanographic parameters. Many databases or data collections that compile information on the Western Mediterranean Sea can be found in literature. Miller (1970) and Guibout (1987) assembled data gathered along selected transects to describe the most important oceanographic features of the Mediterranean Sea. The World Ocean Database (WOD), first released in 1994 and updated approximately every four years, also includes Mediterranean data. The goal of this global project is to stimulate international exchange of modern oceanographic data and encourage the development of regional oceanographic databases as well as the implementation of regional QC procedures. The 2009 WOD series, WOD09 (Boyer et al., 2009), included vertical profiles of temperature, salinity, dissolved oxygen (DO), phosphate, nitrate and nitrite, silicate, chlorophyll-*a*, pH, and alkalinity among other parameters. Products derived from these databases, such as objective analyses of climatologies of variables at standard depths, have also been made available as a separate atlas entitled World Ocean Atlas (WOA). The most important databases exclusively dedicated to the Mediterranean Sea are the MEDAR-MEDATLAS database (MEDAR-Group, 2002) and the MATER database (Monaco & Peruzzi, 2002).

The first climatological analysis of a historical database in the Mediterranean Sea was performed by Brasseur et al. (1996) who reconstructed the three-dimensional temperature and salinity fields. Maillard et al. (2001) computed seasonally and spatially averaged vertical profiles of temperature, salinity, dissolved oxygen, nutrients, and chlorophyll-*a* from *in situ* observations and EU/MTPII/MATER databases under the MEDAR MEDATLAS project.

Oceanographic studies require high-quality data in order to describe temporal and spatial variability of physical, chemical, and biological properties of the oceans. A common problem of databases derived from a variety of methodologies and equipments used for data or samples acquisition and processing arises, leading to a need for establishing an objective QC process to ensure data comparability. A study on the quality of oceanographic data has been carried out by UNESCO/IOC and authors such as Levitus (1982) and Boyer et al. (2009). Levitus (1982) described the QC applied to the Climatological Atlas of the World Ocean (ranges, static stability, statistical, and spurious data checking) and discussed examples of QC application, such as those that allowed legitimatizing the data indicating the presence of Mediterranean water eddies in the Atlantic Ocean or the issue of how often the data must be registered to build really representative average values, useful for defining longterm climatological trends. QC procedures have also been described for the whole oceans (UNESCO et al., 1993; Conkright et al., 2002). Fichaut et al. (1998) defined the ranges of temperature, salinity, and dissolved oxygen in the Mediterranean Sea. The data QC adopted within the data management work of MEDAR-MEDATLAS II was based on specific protocols (Maillard et al., 2001) designed according to international methods and standards (UNESCO et al., 1993). Numerical checks were performed on the incoming data, evaluating their fitting within prescribed ranges (e.g., narrow and broad-range values) and reference statistics. A major difficulty in this process of data validation arises as a marine environment exhibits some trends and/or interannual variations while the existing climatologies, which are representative of the mean and variance of the oceanographic fields, are time-independent (Manca et al., 2004).

In general, existing hydrographic regional databases, including those referring to the Mediterranean Sea, are associated with a set of similar cruises or a wide set of projects. The former present the advantage of a homogeneous control of the data production, its documentation, and processing; but usually their spatio-temporal coverage is insuffcient for detecting mid- or longterm trends. The latter can have both a wider spatial and temporal extension, but because they integrate data from different sources, generated by using diverse equipments, analysis, and post-processing protocols, they usually show a greater dispersion of data values and hence their utility in accurately describing the evolution of hydrographic scenarios is limited. All these considerations are made without underestimating the opportunity and the importance of hydrographic databases already existing in the Mediterranean area.

IBAMar intends to overcome the aforementioned problems, which stem from a lack of spatio-temporal coverage or data heterogeneity, by applying an exhaustive reprocessing of data sets gathered from oceanographic surveys carried out by the Balearic Oceanographic Center of the Instituto Español de Oceanografía (COB-IEO) in the Balearic Sea and the neighboring areas from 1972 onwards within the framework of several successive research projects. Thus, despite a variety of data sources, a close relationship, even identity, among the research teams in charge of the different projects has facilitated a complete standardization of the whole data set. In the first phase, it put together all hydrographic data obtained from vertical casts with multiparametric probes [CTDs (conductivity, temperature, depth) equipped with additional sensors], but it has been recently extended to include biochemical data obtained from water samples taken by means of oceanographic Niskin or Nansen bottles. The result is a standardized database focused on the Balearic region with a main core composed of temperature (T), salinity (S), dissolved oxygen (DO), fluorescence, and turbidity data sets complemented with dissolved inorganic nutrients (phosphate, nitrate, nitrite, and silicate) and chlorophyll-*a* ones.

The Balearic Islands represent a natural limit between the Algerian and the Balearic sub-basins of the Western Mediterranean. As a consequence of their topography, which conditions the exchanges between two sub-basins, there are significant differences between the general hydrodynamic conditions that affect the Northern and Southern regions of the islands. . On the surface, the areas around the Balearic archipelago are characterized by the confluence of the recent Atlantic water (AW) masses entering the Mediterranean through the Strait of Gibraltar and the more saline resident AW, which triggers ocean fronts that affect the hydrodynamics of the whole region (López-Jurado et al., 2008). In this area two intermediate water masses can be detected, the Levantine Intermediate Water (LIW) and the Western Mediterranean Intermediate Water (WIW). The LIW is formed in the Eastern Mediterranean Sea and is characterized by an absolute maximum of salinity, a relative maximum of temperature, and an absolute minimum of dissolved oxygen (DO). The WIW is a winter-mode water, not necessarily present every year, formed seasonally during winter convection processes occurring both in the open sea and over the shelf, mainly in the Northern areas close to the Gulf of Lion but also at lower latitudes (Vargas-Yáñez et al., 2012). The Western Mediterranean Deep Water (WMDW) can be found below a depth of 1000 m, formed during deep winter convection events in the Gulf of Lions and in the Ligurian Sea (MEDOC-Group, 1970). A detailed description of regional water masses and their dynamics can be found in López-Jurado et al. (2008).

The goal of this article is to present the IBAMar database. Sampling methodology, pre-processing of data, and QC protocols are described first. Database organization is briefly explained, and some results on metadata statistics and mean values of different regions are presented in the following sections together with a brief mention of the main scientific results that could be obtained using the IBAMar database. Subsequently, the database policy is defined, and finally some conclusions are presented.

2 MATERIAL AND METHODS

A first version of this database put together only hydrographic data acquired with CTDs equipped with additional sensors during various oceanographic projects developed by the COB-IEO from 1990 to 2010. In a second step, this database was extended to include data from the former and also from biochemical data generated by the analysis of water samples taken using Niskin or Nansen bottles within the same projects. Thus IBAMar (López-Jurado et al., 2014) now collects data from 1974 to 2012, and it is designed to be continually updated with new data. Nowadays, this database brings together data from 164 oceanographic cruises including 7114 sampling stations grouped under 10 major projects. The IBAMar database covers approximately 210 900 km^2 in the Western Mediterranean Sea (see Figure 1), from the Strait of Gibraltar to the east of Menorca and from the Alboran Sea and the Algerian sub-basin to the Gulf of Lions.

2.1 Data acquisition and pre-processing

The total number of sampling stations collected exceeds 7000. In most of these sampling stations (93.6%) hydrographic profiles were obtained by means of CTD probe casts. The main characteristics of CTDs are shown in Table 1, which shows performance of the stations with each CTD model. The CTD model SBE911 was used in most of the cases followed by the SBE25; the CTD model MARK III, mainly used in the 1980s and 1990s, and the old CTD SBE19 that has been improved by the model SBE19plus. IBAMar does not include data from other probes such as mechanical bathythermographs (MBTs) or expendable bathythermographs (XBTs) whose quality data would not be comparable to that of the aforementioned CTDs.

Some CTDs were additionally equipped, from 2001 onwards, with SBE43 sensors for dissolved oxygen determination, Sea Tech, Wetlab (ECO Fluorometer) or Sea Point sensors for fluorescence measurements, and Sea Point sensors for turbidity determination. When the CTDs were deployed integrated in a carousel water sampler, salinity and dissolved oxygen concentrations were calibrated by chemical determination of these variables using *in situ* water samples taken from the hydrographic bottles. Specifically, dissolved oxygen concentration was determinded by the Winkler method (Strickland & Parsons, 1972), and salinity was determined using an Autosal salinometer. Calibrations were performed at selected depths in the water column at least once per cruise when a campaign was shorter than a week and at least in the beginning of a campaign and the end when it was of a longer duration. During TUNIBAL campaigns, calibrations were done every three days. When these *in situ* water samples were not available, data from the CTD sensor were compared with historical calibrations of the same sensor. In addition, the SBE25 was, in some campaigns, cross-calibrated with a SBE911 attaching it to the same carousel water sampler at least once every campaign when both CTDs were embarked. Generally, once every two years, all the CTD probes and additional sensors were factory calibrated.

All data from Seabird (SBE) CTDs were reviewed and pre-processed with the software provided by Sea-Bird Electronics, Inc. Details of this software can be found in the Seabird data processing manual, (see http://www.seabird.com/). Data for each station were processed at the original sampling frequency and averaged every 1 dbar. The sequence of routines used, their coefficients, and the general sampling methodology are described in the RADMED Protocols (see http://www.repositorio.ieo.es/e-ieo/). Only a small fraction of the sampling stations (6.4 %) were obtained exclusively by means of hydrocasts where only Nansen or Niskin bottles were used.

In all cases, water samples from hydrographic bottles for nutrient analysis were kept frozen until their determination in the laboratory. Nitrate, nitrite, and silicate concentrations were measured according to the method of Armstrong et al. (1967), modified by Grasshof (1969). Phosphate concentrations were determined by the method of Treguer and Le Corre (1975). All these methods were adapted to the Mediterranean oligotrophic waters. Analyses were performed with a Technicon Autoanalyzer AAII. The values of chlorophyll-*a* correspond to valuations based on fixed volumes of filtered water carried out using spectro-fluorimeter measurements by fluorimetry (Holm-Hansen et al., 1965).

2.2 Quality control protocols

Different technologies and methodologies by independent teams were used during the four decades of oceanographic data sampling carried out by the COB-IEO. However, in order to minimize possible biases or errors due to pre- or post-processing, different QC protocols have been applied to each variable, checking all the gathered data exhaustively before their inclusion in the IBAMar database.

QC for IBAMar data includes both manual and automatic procedures, which were performed on each vertical profile and also on profiles grouped by areas.

Manual procedures include visual check of metadata, duplicate data, units, and obvious outliers. Automatic procedures include a variable range check per area and depth, by season in some cases, and also a check for spikes and density inversions. Spikes in salinity and temperature measurements can be detected in some stations due to steep gradients of these values at some depth, e.g., the thermocline. A spike is detected by comparing a data value at a depth, $V2$, with the value at the previous depth, $V1$, and the value at the next depth, $V3$ (Maillard et al., 2001).

$$\left|V2-\frac{(V3+V1)}{2}\right|-\left|\frac{(V1-V3)}{2}\right|>TH,$$

TH being a threshold value settled to 2 ºC for temperature and 0.3 (PSU) for salinity.

The check for inversions calculates the potential density, σ_Θ, at each depth *V2*. If σ_Θ at *V3* plus a noise value is greater than σ_Θ at *V2*, then the stratification is stable. Otherwise, the stratification is unstable. A noise level of 0.05 kg m^{-3} is defined for data between 0 and 100 dbar and 0.03 kg m^{-3} for data from 100 dbar and below.

Three different regions, defined by their different hydrographic features, have been considered for the QC analysis: the North of the Balearic Islands (NBI) includes the Balearic Sea and the Gulf of Lions and the South of the Balearic Islands (SBI) includes the Algerian sub-basin and the Alboran Sea (AS). Coastal areas, defined as areas of shallow water close to the coast, and open sea areas have also been considered separately. The limit between the coastal and the open sea is defined to be in the 200 m bathymetric line for convenience. The different regions and the 200 m isobaths are shown in Figure 2.

The vertical profiles were divided into 27 standard levels (SLs), which correspond to the first 27 standard levels defined in the Climatological Atlas of the World Ocean (Levitus et al., 1998). Maximum and minimum depth of each SL are shown in Table 2. Data values within each SL were averaged in order to perform statistical analysis and to construct the vertical profiles included in the freely available version of the IBAMar. The standard deviation (SD) of the data values was calculated within each zone and the SL to define narrow and broad range bands set as 3 SD and 5 SD, respectively. The broad band was used for coastal areas and the upper SL due to their high variability.

A quality flag is assigned as an integer number to each data value depending on the result of the QC checks as shown below in Table 3.
1 – CORRECT VALUE means that the data mean value is within the narrow range of the SD check and the other checks are positive,
2 – VALUE INCONSISTENT WITH STATISTICS means that the data mean value is outside the narrow range of the SD check,
3 – SPIKE means that the data has been flagged as a spike by the method described above,
4 – DENSITY INSTABILITY means that the data present a density inversion at the corresponding depth,
5 – not used,
6 – FALSE VALUE means that the data mean value is outside the broad range of the SD check. This is used only for nutrients and chlorophyll-*a*.

Thus, data were flagged with number 2 if they were outside an envelope of 5 SD from the mean in coastal or above 50 m depth (SL 1, 2, 3, or 4); whereas an envelope of 3 SD has been applied to the open sea and below 50 m depth data. In special areas like ports, delta rivers, closed areas, etc., a different definition of narrow and broad ranges was used because sometimes it is necessary to include unusual but legitimate data. The mean values and SD for nutrients and chlorophyll-*a* were calculated seasonally owing to their strong seasonal variability.

3 DATABASE ORGANIZATION

IBAMar database groups the oceanographic stations by cruises, which are organized by projects. The stations that are made individually outside a regular cruise are grouped as virtual “cruise”, whose name depends on the ship, month, and place. Occasionally, a single station is sampled by different CTDs at the same tim,e and the CTD type field can differentiate these data. The database resulting from this process consists of 7114 stations carried out during 164 cruises grouped into 10 major projects as summarized in Table 4.

Each oceanographic station includes a header with metadata and a matrix with data values. Station headers contain the cruise name, research vessel, station number, date, time, location, CTD type or Bottle, and the bottom depth. The matrix of data includes the data values of the different parameters organized by columns. QC flags for each value (by dbar or SL depending on the IBAMar version) are also included in the matrix of data. Pressure, salinity, and temperature data columns are mandatory while the rest are optional. Depending on the cruise, the station profile may also have discrete data at the depth where the sample bottles were closed. These discrete data include nitrate, nitrite, phosphate, silicate, and chlorophyll-*a* values.

IBAMar can be accessed using the Ocean Data View (ODV) software package (Schlitzer, 2007), which allows an interactive exploration, analysis, and visualization of the spatial and temporal distribution of data and the visualizations of the profiles.

4 STATISTICS

4.1 Metadata statistics

The number of measurements for all parameters is shown in Figure 3. Note the different scale for CTD data (left) and for discrete data (right). Salinity and temperature are the parameters that have a larger number of determinations while chlorophyll-*a* is one that has less.

Figure 4 shows the number of times each SL was sampled for any parameter included in the IBAMar database. The upper meters were sampled almost 7000 times. This number decreases as the depth increases because there are fewer stations with higher depths and also because some projects only sample the upper ocean layers. For example, in the campaigns carried out under the BALEARIC TUNA program, CTD casts only reached 350 or 650 m depth (Alemany et al., 2010).

As mentioned earlier, the spatial location of IBAMar database ranges from the Strait of Gibraltar to the east of Menorca and from Alboran and Algerian sub-basin to the Gulf of Lions. Sampling stations are distributed in the zones defined above (see Figure 2). The bulk of IBAMar stations are within zone NBI (55.2%) and zone SBI (37.6%) with only 8.0% of stations within the AS zone. Coastal stations with bottom depth less than 200 m represent 41.0% of the data as compared to oceanic stations that comprise 59.0% of the database.

The seasonal distribution of sampling stations is shown in Figure 5. Summer (in red) is the most extensively sampled season. This fact is related to the type of studies usually carried out by the COB-IEO, which give priority to hydrographic projects on variability and circulation of water masses and multidisciplinary projects studying the spawning conditions of different fish stocks. Moreover, the bad winter weather (unfavorable for cruises) besides conducting maintenance tasks of research vessels in winter also explains the relative scarcity of data from winter campaigns (in black).

Figure 5 also shows the distribution of sampling stations through the years. Some years were more extensively sampled, with more than 500 stations, e.g., 2004, 2003, 2005, and 2009 with 555, 575, 583, and 716 sampling stations respectively. Years with fewer than 50 stations are 1982 and 1980, which include 25 and 49 sampling stations respectively.

Table 1 shows the number of casts made with different CTD type and bottles, accuracy in measurements, and sampling frequency. A high number of CTD stations (93.6%) can be observed where the SBE911 model was the predominant probe (58.0% of the total).

The distribution of sampling stations by major project (Table 4) shows that RADMED, BALEARIC TUNA, and CIRBAL, all of them over regular grids, are the projects with the greatest number of sampling stations (19.6, 19.9, and 26.7% respectively).

4.2 Quality control results

After QC procedures, every set of data was flagged with values as shown in Table 3. Flag number 1 represents 98.94% for salinity, 98.96% for temperature, 99.58% for dissolved oxygen, 98.46% for turbidity, 98.51% for fluorescence, 98.76% for nitrate, 98.32% for nitrite, 99.03% for phosphate, 98.68% for silicate, and 98.51% for chlorophyll-*a*. Flags different from 1 do not necessary mean “bad” data, but data that have to be used carefully. Flag number 2 represents 1.06% for salinity, 1.04% for temperature, 0.42% for dissolved oxygen, 1.54% for turbidity, 1.49% for fluorescence, 1.05% for nitrate, 1.28% for nitrite, 0.86% for phosphate, 1.24% for silicate, and 1.30% for chlorophyll-*a*. Flag number 3 is less than 0.001% for salinity and temperature. Flag number 4 is 0% for salinity and temperature. Flag number 6 represents 0.18% for nitrate, 0.39% for nitrite, 0.12% for phosphate, 0.08% for silicate, and 0.19% for chlorophyll-*a*.

Table 5 summarizes the properties of the water masses at depths corresponding to their core by zone, indicating their mean values together with their SD and the number of times they have been sampled. The high variability observed for nutrients and chlorophyll-*a* is because the averages are calculated independent of the season. In a stratified ocean there is a strong difference in the range of nutrients and chlorophyll-*a* between winter and the rest of the year. For that reason, the QC flag number 6 has been specially analyzed by season to avoid the issue of winter ranges masking the values for the rest of the year.

Vertical profiles for averaged values, constructed for all physical and biochemical parameters compiled in the IBAMar database for each of the six zones and SL, are shown in Figure 6 (coast and open sea down to 200 m) and in Figure 7 (open sea down to the bottom).

The profiles were calculated with all data available independent of the flags assigned to each one. The Western Mediterranean seasonality is critical for explaining the variability of both physical and biochemical properties in the water column (Manca et al., 2004). The principal events related to seasonality are convective movements in winter, stratification in summer, and biological activity such as spring phytoplankton blooms and river run-off. Other events such as upwelling and anthropogenic effects on coastal areas are also to be taken into account. It should be noted that the distribution of water masses is different between zones, seasons, and years. The AS is characterized as the receiving basin of surface AW and also the end of the cyclonic circulation path of the LIW and the WMDW that are part of the outflow of the Mediterranean in the Gibraltar Strait as Mediterranean Outflow Water (MOW). The SBI is an area of confluence between the fresh AW and the saltier resident AW, where the intermediate and deep waters also arrive from their place of origin after a long trajectory from the Northwestern Mediterranean. The NBI contains surface AW that has stayed longer in the Mediterranean and is where the winter convection processes lead to the formation of intermediate and deep waters. From IBAMar data it is possible to identify the water masses affected by these processes and to monitor their cores through the minima of temperature, maxima of salinity, and minima of DO that characterize the water masses in the different zones.

Chlorophyll-*a* measurements are conducted in the euphotic layer from discrete depths considering that these values are negligible below 250 m. The distribution of chlorophyll-*a* and phytoplankton production, in both open and coastal waters, is directly linked with features such as stratification and mixing in the water column, thickness of the photic layer, and nutrient availability. Winter mix and exchange with the Atlantic Ocean happen on a broad scale while upwelling or river discharges act locally. Fluorescence is used as an indicator of chlorophyll concentration in the sea and provides a method for determining the biological activity in the water column. The main profile for chlorophyll-*a* and fluorescence (see Figures 6 and 7) are similar, showing differences between zones. In SBI and NBI (both coast and open sea) fluorescence, the mean values (or chlorophyll-*a*) reflect the summer stratification with a deep maximum (around 75 dbar) in the limit of the photic layer where nutrients are available for photosynthesis. In AS (coast and open) the mean values show that the maximum is closer to the surface due to the upwelling, which provides nutrients to the upper layers. The DO at the surface and the mixed layer are at equilibrium with the atmosphere. The observed increase below the surface is the result of the biological activity of phytoplankton, which decreases under the photic zone as it is then consumed by biological degradation or remineralization of organic matter. The DO profiles are a mirror image of the nutrient (silicate, nitrate, phosphate) profiles as shown in Figure 6. The minimum of DO is observed at the LIW depth, in the Alboran Sea, in agreement with previous studies (Packard et al., 1988).

Turbidity values through the water column in open zones are small and lower than 0.2 FTU, the increase in coastal zones reaching over 0.40 FTU as shown in Figure 6. This increase in turbidity is due to the proximity of the sea floor, the action of storms, convective mixing, or the presence of bottom currents that generate turbulent bottom friction with consequent increase of resuspended matter. Turbidity in the AS coast is higher than in the rest of the zones (between 0.27 ± 0.16 FTU in SL 9, 151–200 m, and 0.46 ± 0.58 FTU in SL 2,11–20 m). This is due to the high current and upwelling events in this area of the Mediterranean Sea.

Silicate, nitrate, and phosphate (see Figure 6) follow the general pattern of the oligotrophic Mediterranean waters. They are depleted in the photic zone, which is generally attributed to phytoplankton consumption and the scarce supply of nutrients from the deep layer. Their concentrations in the mean profiles increase from the surface until SL 16 (701–800 m) and then remain constant down to the sea floor; this is due to the release of nutrients during the respiration of organic matter. Nitrite is released by phytoplankton during its growth, and its concentrations are very low in comparison with nitrate. In the AS coastal zone, nutrient concentrations are always slightly higher than in the NBI and SBI coastal zones, in general due to the upwelling in coastal areas and also due to the inputs of nutrients from the Atlantic Ocean, which fuel the photic zone for the phytoplankton growth, as shown for the whole column in open sea (see Figure 7).

4.3 Use of IBAMar database as a reference for further scientific studies

IBAMAR features show that it is a regional and fairly homogeneous database with regard to the type of data available. A very high percentage of its data correspond to CTDs with the same type of sensors and having received a similar post-process. In addition, the main core of cruises use a regular grid of stations so that the same station or section is repeated many times over the years. This fact favors the development of studies related to the annual cycles of different variables, their seasonal and interannual variability, the effects of winter convective processes, the presence of water masses, mesoscalar structures, transport and exchange between basins, cycles, trends, and possible climate changes as well as environmental and ecological studies of species.

In this context, and among other studies, IBAMar has allowed the monitoring of the evolution of the thermohaline anomaly of the deep waters in the West Mediterranean (WMED), first observed in 2005 (López-Jurado et al., 2005) and which now affects the whole Western Mediterranean Basin.

The IBAMar database has been used to decribe the environmental differences between fishing grounds north and south of the islands (López-Jurado et al., 2008). Using IBAMar data it was possible to develop a climate index, the IDEA index, which relates anomalies of air temperature during winter in the Gulf of Lions with the presence of WIW in the Balearic channels (Monserrat et al., 2008).

IBAMar data have also been used to characterize the spatial distribution and the temporal variability of the dissolved oxygen around the Balearic Islands (Balbín et al., 2014). In that work, seasonal climatology of temperature, salinity, and DO were computed at selected stations of the RADMED monitoring program. These climatologies reflect the appearance of WIW around the Balearic Islands in winter, spring, and summer that the MEDAR climatologies do not show. The use of BALEARIC TUNA data and other data from the IBAMar database reveals that the hydrographic early summer scenarios around the Balearic Islands are related to the winter atmospheric forcing in the Northwestern Mediterranean Sea (Balbín et al., 2014b).

The MEDATLAS and IBAMar databases were used to construct the longest temperature and salinity time series ever analyzed in the Western Mediterranean (1900 to 2008) (Vargas-Yáñez et al., 2010). These time series show that both the upper and intermediate layers have warmed throughout the twentieth century. Longterm and decadal variability in the upper layer correlate with surface air temperature in the northern hemisphere and heat absorbed by the upper North Atlantic Ocean, suggesting that the time series analyzed in that work reflect the present heat absorption of the oceans in the context of global warming.

These examples illustrate the main advantages of the IBAMar database sampled under different monitoring programs over regular grids (CANALES, CIRBAL, BALEARIC TUNA, RADMED). This supports the historical analysis of data, such as the water masses transport in the Balearic channels (Heslop et al. 2012) or the ecosystem studies around the Balearic Islands (Balbín et al. 2014; Alemany et al., 2010; Reglero et al., 2012; Hidalgo et al., 2014; Torres et al., 2011; Rodríguez et al., 2013).

4.4 Comparison of IBAMar with other databases

MEDAR (MEDAR-Group, 2002) is a database that has inspired the structure and organization of the IBAMar database. The number of CTD data within the IBAMar area included in MEDAR is of the order of 6000 and extends from the 1970s to 2000. After that year, there is no upgrade in this database to include all posterior samples.

The World Ocean Database (WOD13, Boyer et al., 2013) includes CTD data from the 1970s to 2012, but within the IBAMar area, there are very few CTD data after year 2000 probably because most of the data are mainly from the MEDAR database and any posterior upgrade depends on a voluntary decision of the data producers.

The CORIOLIS data portal (http://www.coriolis.eu.org/) provides a single access point for *in situ* data and products. It receives data from Argo, French research ships, voluntary observations, merchants' ships, moorings, and the WOD (not in real time for the last one and only for CTD). The available data are taken from different types of instruments, mainly Argo floats, XBT, CTD, XCTD, and moorings. With regard to CTD data within the IBAMar area, CORIOLIS contains mainly WOD data and also data from some French research ships working mainly in the Gulf of Lions.

SeaDataNet (http://www.seadatanet.org/) is an international project that aims to provide a web service to enable access to historical datasets owned by national data centers. Users of SeaDataNet who want to retrieve datasets from multiple data centers send their request, which is analyzed and split into requests to the data centers concerned. Finally, the user receives the authorization to retrieve the data. It is not a database but a website connecting data centers. The recently initiated EMODnet Physics web site (http://www.emodnet-physics.eu/portal/) provides access to physical data and metadata. Near real-time data and metadata are provided by data owners. The last 60 days are freely viewable and downloadable while access to older data (monthly archives) requests credentials. These are websites connecting users with data providers, but they do not work as real databases.

5 DATA AVAILABILITY AND DATA POLICY

All IBAMar products including data are subject to the following conditions:

1..IBAMar metadata and station values at SL are freely accessible from the IEO website http://www.ba.ieo.es/ibamar/, but the origin of the data must always be explicitly acknowledged by citing López-Jurado et al. (2014) and this work.
2. Detailed data from IBAMar (station values every dbar) are available under request to the IBAMar database coordinator through the IEO website. The application should include a description of the intended use of the data and an explicit agreement with IBAMar data policy.
3. Data will be provided under a collaboration agreement between the requesting institution and the IEO. This agreement has to fulfill the data policy of the IBAMar database and also the data policy of the specific programs that fund the requested data acquisition.
4. As IBAMar collects data from different projects, any user who needs help in making use of the data and its interpretation might consider the project coordinator collaboration for co-analysis of data and, in this case, co-authorship of the published results. In any case, users must always acknowledge data sources and cite them as agreed with the project coordinator of the requested data.
5. Data users should not give to third parties any IBAMar data or product without prior consent from the data owner.
6. Data users must respect all restrictions on the use or reproduction of data. The use or reproduction of data for commercial purposes might require prior written permission from the data source.

6 CONCLUSIONS

The IBAMar database has been gathering and preserving oceanographic data from waters around the Balearic Islands in the Western Mediterranean Sea from the 1970s until the present. Most of the hydrographic data were obtained with similar multiparametric probes (CTDs) with the same post-processing protocols. Many of the data are sampled under different monitoring programs over regular grids. Therefore, it provides a regional database of homogeneous, good quality data that is robust enough to generate a statistical background that will permit improvement in the knowledge of the waters around the Balearic Islands. The IBAMar database enables the archiving of data obtained during projects developed by COB-IEO with proper management and documentation. Merging IBAMar with MEDAR in the Balearic Sea will help to better calculate climatologies of the different parameters, to characterize time series at critical points of the sea-like channels and deep water formation areas, to study oscillations and trends in the oceanographic parameters, which are important for climate change studies and the study of its environmental impact. IBAMar also supports biological and ecosystem research in these waters. The database described in this work will be an important and useful tool for future research projects and studies to be developed in the Western Mediterranean Sea.

7 ACKNOWLEDGEMENTS

The oceanographic database is the basis of this work. Exploring the sea and obtaining data are not easy. For this reason we want to acknowledge those who were involved in this task, the crew and the research staff. We also thank all the those who worked behind the scenes in the analysis laboratories. IBAMar was created mainly to preserve data and generate a background of regional information that will serve to develop new projects and knowledge. Therefore, we want to thank all those who have understood the message that "in general, non-public information is ephemeral and leaves no trace". Thanks to all the colleagues who participated in the initial monitoring projects developed in the different IEO laboratories without external funding whose data configured the nucleus of IBAMar (ECO-Málaga, ECO-Murcia, ECO-Baleares, CANALES, CIRBAL, and RADMED) and thanks to all the colleagues from other projects who have generously contributed their data and whose traces can be found in IBAMar (TUNIBAL, TUNABIT, BLUEFIN, EFLUBIO, ATAME, IDEA, IDEADOS, etc.). We also acknowledge E. Heslop (SOCIB) and F. Alemany for their invaluable comments on this manuscript.

IBAMar was funded by the "Consejería de Innovación, Interior y Justicia (DG de Investigación, Desarrollo Tecnológico e Innovación)" of the "Govern de les Illes Balears" under projects AAEE0022/08 and AAEE0054/09

8 REFERENCES

Alemany, F., Quintanilla, L., Velez-Belchí, P., García, A., Cortés, D., Rodríguez, J. M., Fernández de Puelles, M. L., González-Pola, C., & López-Jurado, J. L. (2010) Characterization of the spawning habitat of Atlantic bluefin tuna and related species in the Balearic Sea (western Mediterranean). *Progress in Oceanography 86*, pp 21–38.

Armstrong, F., Stearns, C., & Strickland, J. (1967) The measurement of upwelling and subsequent biological process by means of the technicon autoanalyzer® and associated equipment. In *Deep Sea Research and Oceanographic Abstracts 14*, pp 381–389. Elsevier.

Balbín, R., López-Jurado, J. L., Aparicio-González, A., & Serra, M. (2014a) Seasonal and interannual variability of dissolved oxygen around the balearic islands from hydrographic data. *Journal of Marine Systems 138*, pp 51-62.

Balbín, R., López-Jurado, J. L., Flexas, M. M., Reglero, P., Vélez-Velchí, P., González-Pola, C., Rodríguez, J. M., García, A., & Alemany, F. (2014b) Interannual variability of the early summer circulation around the Balearic Islands: driving factors and potential effects on the marine ecosystem. *Journal of Marine Systems 138*, pp 70-81.

Boyer, T. P., Antonov, J. I., Baranova, O. K., Garcia, H. E., Johnson, D. R., Locarnini, R., Mishonov, A., O'Brien, T., Seidov, D., Smolyar, I., & Zweng, M. M. (2009) World Ocean Database 2009. NESDIS 66, NOAA Atlas.

Brasseur, P., Beckers, J. M., Brankart, J. M., & Schoenauen, R. (1996) Seasonal temperature and salinity fields in the mediterranean sea: Climatological analyses of a historical data set. *Deep Sea Research Part I: Oceanographic Research Papers 43*(2), pp 159–192.

Conkright, M., Garcia, H., O'brien, T., Locarnini, R., Boyer, T., Stephens, C., & Antonov, J. (2002) *World Ocean Atlas 2001. Volume 4, Nutrients*. National Oceanic and Atmospheric Administration.

Fichaut, M., Balopoulos, E., Dooley, H., Garcia-Fernandez, M., Iona, A., Jourdan, D., Baudet, L., & Maillard, C. (1998) A common protocol to assemble a coherent database from distributed heterogeneous data sets: The medatlas database experience. *Marine science and technology programme: Experiences in project data management,* Office for Official Publications of the European Communities, Luxembourg, pp 311–327.

Grasshof, K. (1969) *On an Apparatus for Similtaneous Determination of 6 Chemical Compounds in Sea Water with Digital and Analogue Output 20*, Verlag: Hamburg, Germany.

Guibout, P. (1987) Atlas Hydrographique de la Mediterranée. Technical report, IFREMER.

Heslop, E., Ruiz, S., Allen, J., López-Jurado, J.L., Renault, L., & Tintoré, J. (2012) Autonomous underwater gliders monitoring variability at "choke points" in our ocean system: A case study in the Western Mediterranean Sea. *Geophysical Research Letters 39*, L2060424, doi: 10.1029/2012GL053717.

Hidalgo, M., Reglero, P., Álvarez-Berastegui, D., Pérez-Torres, A., Álvarez, I., Rodríguez, J. M., Carbonel, A., Zaragoza, N., Tor, A., Goñi, R., M. S., Babín, R., & Alemany, F. (2014) Hydrographic and biological components of seascape structure the meroplankton community in a frontal system. *Marine Ecology Progress Series 505*, pp 65–80.

Holm-Hansen, O., Lorenzen, C. J., Holmes, R. W., and Strickland, J. D. (1965) Fluorometric determination of chlorophyll. *Journal du Conseil 30*(1), pp 3–15.

Levitus, S. (1982) *Climatological atlas of the world ocean*. US Dept. of Commerce, National Oceanic and Atmospheric Administration (Rockville, Md.).

Levitus, S., Boyer, T., Conkright, M., OÂ'Brien, T., Antonov, J., Stephens, C., Stathoplos, L., Johnson, D., & Gelfeld, R. (1998) *World Ocean Database 1998: vol. 1: Introduction*. Technical Report NESDIS 18, NOAA, Washington DC.

Lopez-Jurado, J.L., Aparicio-González, A., Babín, R., Alonso, J., Amengual, B., Jansá, J., García-Martínez, M., Moya, F., Serra, M., & Vargas-Yáñez, M. (2014) IBAMar DATABASE: 4 decades sampling on the Western Mediterranean Sea. Instituto Español de Oceanografía doi.pangaea.de/10.1594/PANGAEA.831923 and http://www.ba.ieo.es/ibamar/

López-Jurado, J. L., González-Pola, C., & Vélez-Belchí, P. (2005) Observation of an abrupt disruption of the long-term warming trend at the Balearic Sea, Western Mediterranean Sea, in summer 2005. *Geophysical Research Letters 32*(24).

López-Jurado, J. L., Marcos, M., & Monserrat, S. (2008). Hydrographic conditions affecting two fishing grounds of Mallorca Island (Western Mediterranean): during the IDEA Project (2003-2004). *Journal of Marine Systems 71*(3-4), pp 303–315.

Maillard, C., Fichaut, M., & MEDAR/MEDATLAS-Group (2001) *Medar-Medatlas Protocol (V3) Part I: Exchange format and quality checks for observed profiles*. Technical report, IFREMER.

Manca, B., Burca, M., Giorgetti, A., Coatanoan, C., Garcia, M., & Iona, A. (2004) Physical and biochemical averaged vertical profiles in the Mediterranean regions: an important tool to trace the climatology of water masses and to validate incoming data from operational oceanography. *Journal of Marine Systems 48*(1-4), pp 83–116.

Massutí, E., Moranta, J., Gordon, J. D., Morales-Nin, B., & Oliver, P. (2008) The wrapping up of the IDEA project: International Workshop on Environment, Demersal Resources and Fisheries. *Journal of Marine Systems 71*(3–4), pp 221 – 222.

MEDAR-Group (2002) *MEDATLAS 2002 Mediterranean and Black Sea database of temperature, salinity and biochemical parameters climatological atlas*. In *4 CD-ROM*. European Commission Marine Science and Technology Programme (MAST), IFREMER.

MEDOC-Group (1970) Observation of formation of deep water in the Mediterranean Sea. *Nature* 227(1037), pp 1040.

Miller, A. (1970) *Mediterranean Sea Atlas of Temperature, Salinity, and Oxygen. Profiles and Data from Cruises of RV Atlantis and RV Chain*. Technical report, Woods Hole Oceanographic Institution, Massachusetts.

Monaco, A. & Peruzzi, S. (2002) The Mediterranean Targeted Project MATER –a multiscale approach of the variability of a marine system– overview. *Journal of Marine Systems 33,* pp 3–21.

Monserrat, S., López-Jurado, J. L., & Marcos, M. (2008) A mesoscale index to describe the regional circulation around the balearic islands. *Journal of Marine Systems 71*(3-4), pp 413–420.

Packard, T., Minas, H., Coste, B., Martinez, R., Bonin, M., Gostan, J., Garfield, P., Christensen, J., Dortch, Q., Minas, M., et al. (1988) Formation of the Alboran oxygen minimum zone. *Deep Sea Research Part A. Oceanographic Research Papers 35*(7), pp 1111–1118.

Reglero, P., Ciannelli, L., Alvarez-Berastegui, D., Balbín, R., López-Jurado, J., & Alemany, F. (2012) Geographically and environmentally driven spawning distributions of tuna species in the western Mediterranean Sea. *Marine Ecology Progress Series 463*, pp 273–284.

Rodríguez, J., Álvarez, I., López-Jurado, J. L., García, A., Balbín, R., Álvarez-Berastegui, D., Torres, A., & Alemany, F. (2013) Environmental forcing and the larval fish community in the Balearic region (Western Mediterranean) in early summer 2005. Deep-Sea Research Part I : *Oceanographic Research Papers 77*, pp 11–22.

Schlitzer, R. (2007) Ocean data view. Retrieved from the World Wide Web January 4, 2015: http://odv.awi.de/.

Strickland, J. & Parsons, T. (1972) A practical handbook of seawater analysis. *Bull. Fish. Res. Board Can. 167*.

Torres, A., Dos Santos, A., Balbín, R., Alemany, F., Massutí, E., & Reglero, P. (2014) Decapod crustacean larval communities in the Balearic Sea (western Mediterranean): seasonal composition, horizontal and vertical distribution patterns. *Journal of Marine Systems 138*, pp 112–126.

Treguer, P. & Le Corre, P. (1975) *Manuel d'analyse des sels nutritifs dans l'eau de mer. Utilisation del'AutoAnalyser II Technicon.* Occidentale, Univ. Bretagne, Laboratoire de Chimie Marine, Brest, France.

UNESCO, CEC/DG-XII, MAST, & IOC/IODE (1993) Manual of Quality Control Procedures for Validation of Oceanographic Data. In *IOC Manual and Guides 26*. UNESCO.

Vargas-Yáñez, M., Moya, F., García-Martínez, M., Tel, E., Zunino, P., Plaza, F., Salat, J., Pascual, J., Lopez-Jurado, J., & Serra, M. (2010) Climate change in the Western Mediterranean sea 1900–2008. *Journal of Marine Systems 82*(3), pp 171–176.

Vargas-Yáñez, M., Zunino, P., Schroeder, K., López-Jurado, J., Plaza, F., Serra, M., Castro, C., García-Martínez, M., Moya, F., & Salat, J. (2012) Extreme Western Intermediate Water formation in winter 2010. *Journal of Marine Systems,105-108*(0), pp 52–59.

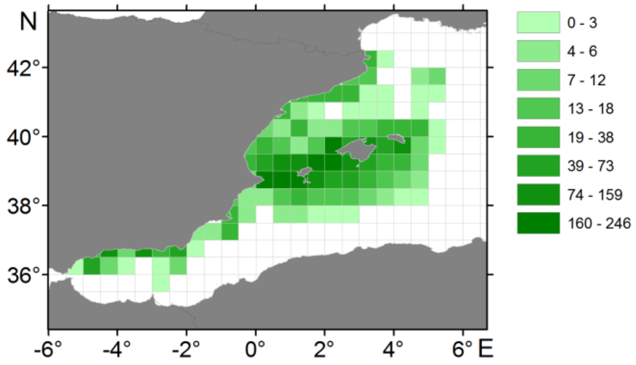

Figure 1. Geographic location of oceanographic stations included in IBAMar. Color scale indicates the number of stations per 1000 km^2 portrayed in a 0.5° x 0.5° grid squares.

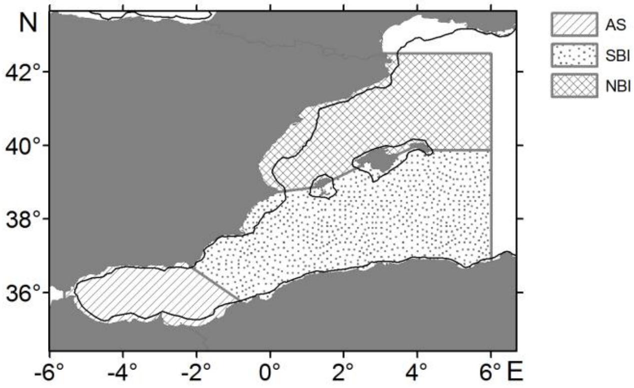

Figure 2. Map of the Western Mediterranean Sea showing the three zones and their geographical limits. AS (Alboran Sea), SBI (South of Balearic Islands), NBI (North of Balearic Islands). Black thick line is the 200 m isobath.

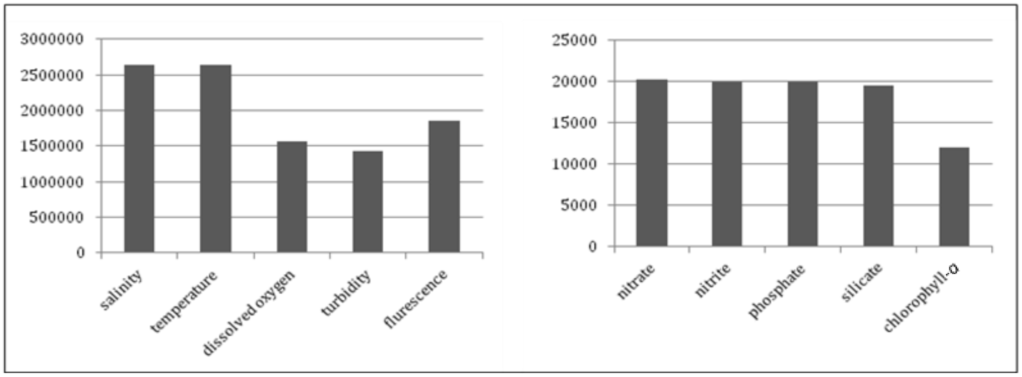

Figure 3. Number of measurements for all parameters within IBAMar

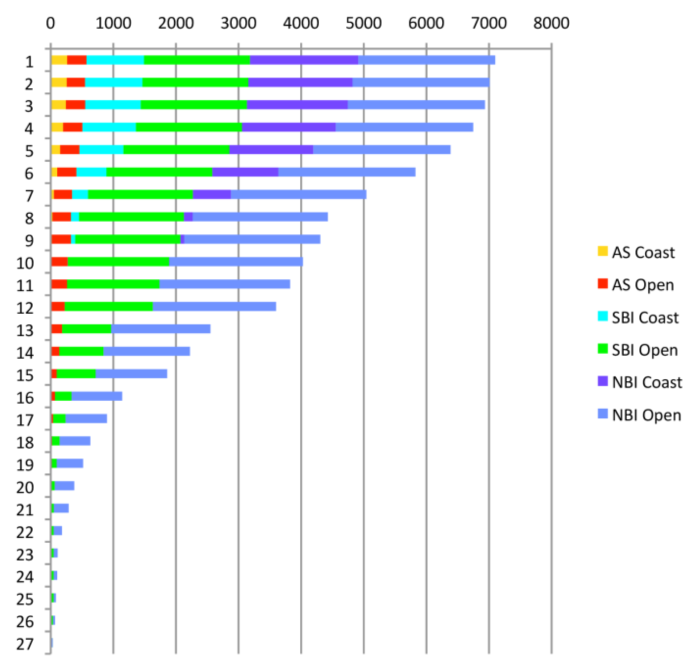

Figure 4: Number of times that each standard level (SL) was sampled by zone

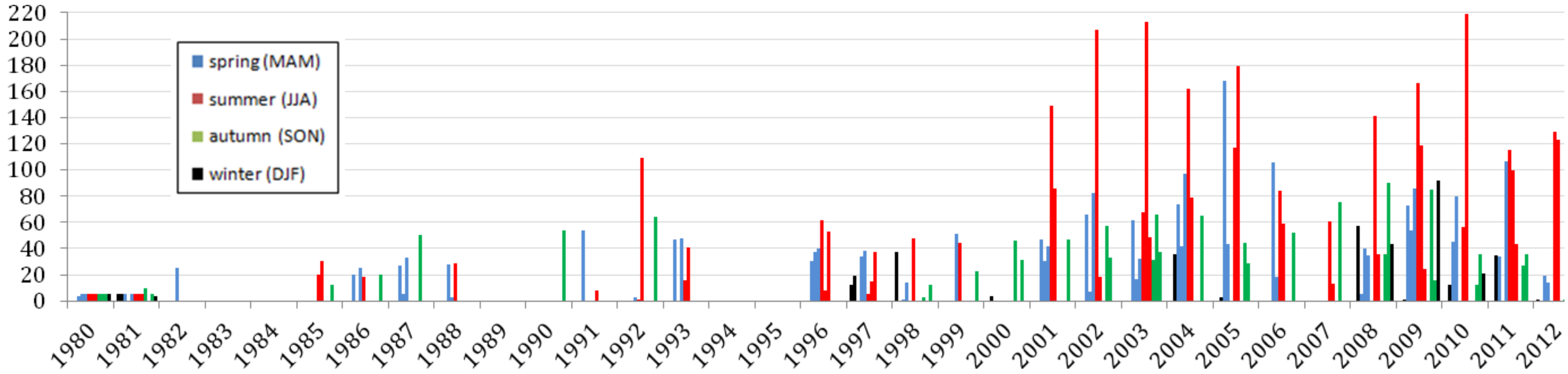

Figure 5. Seasonal distribution of sampling stations

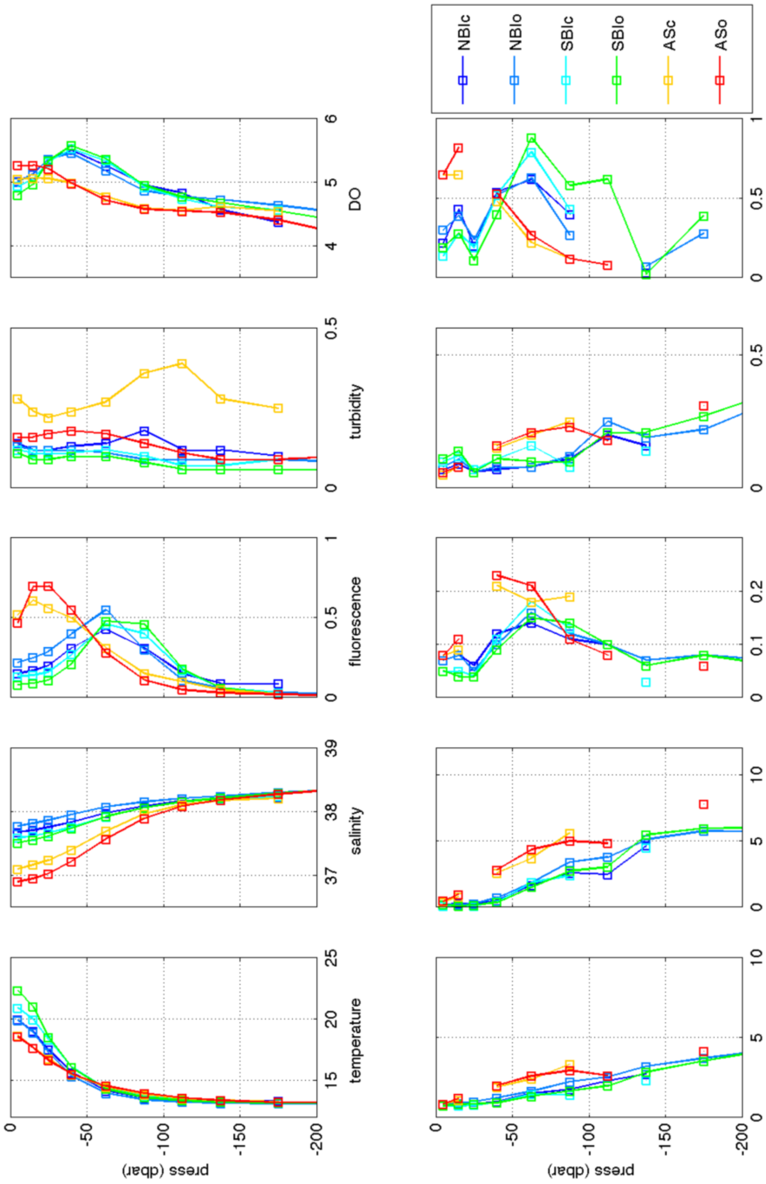

Figure 6. Mean profiles to 200 m depth. Units are as indicated in Table 5.

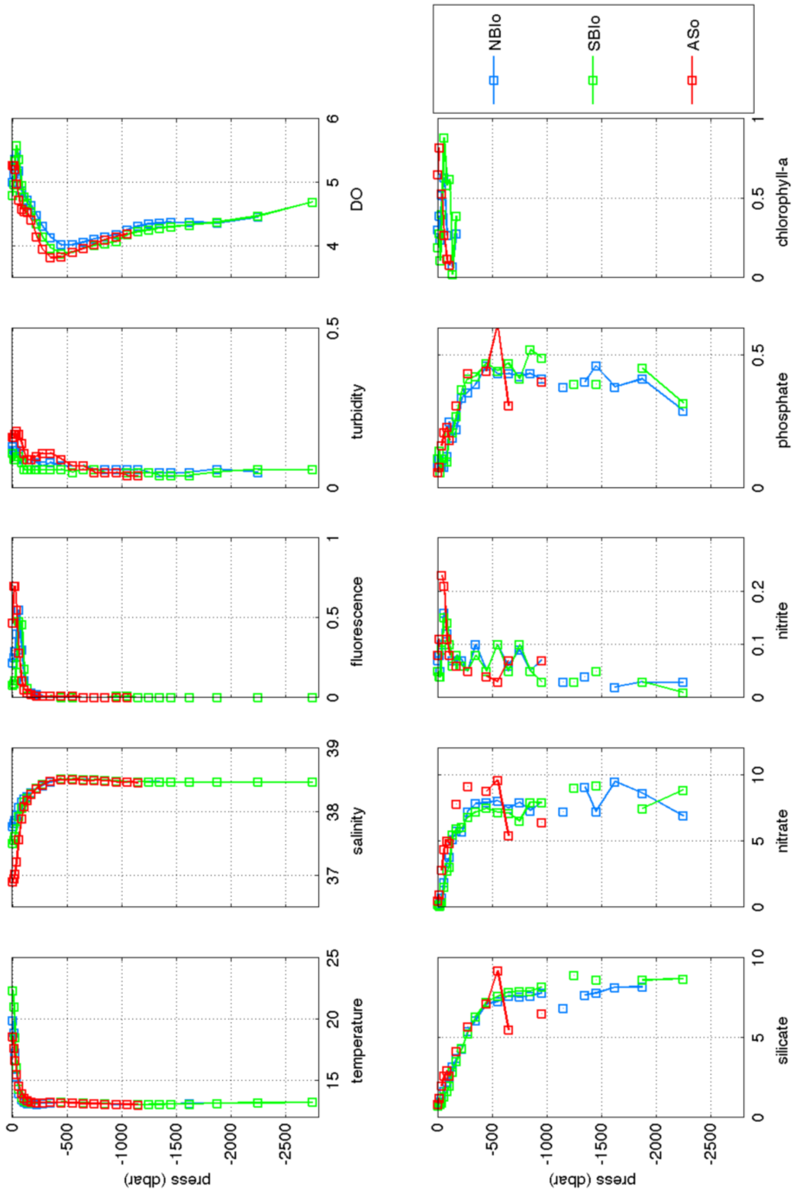

Figure 7. Mean profiles for open sea stations down to the sea floor. Units are as indicated in Table 5.

Table 1. Number of casts made with CTD and bottles in addition to the main features of CTDs

CTD Type	MARK III	SBE911	SBE25	SBE19plus	SBE19	Bottles
Number of casts	167	3868	2016	573	36	454
Sampling frequency (Hz)	31.25	24	8	4	1	–
Temperature accuracy (ºC)	10^{-3}	10^{-3}	10^{-3}	$5\cdot10^{-3}$	10^{-2}	–
Conductivity accuracy (S/m)	$5\cdot10^{-4}$	10^{-4}	10^{-4}	$5\cdot10^{-4}$	10^{-3}	–

Table 2. Minimum and maximum limits (m) of standard levels (SL)

SL	Min. Depth	Max. Depth
1	1	10
2	11	20
3	21	30
4	31	50
5	51	75
6	76	100
7	101	125
8	126	150
9	151	200
10	201	250
11	251	300
12	301	400
13	401	500
14	501	600
15	601	700
16	701	800
17	801	900
18	901	1000
19	1001	1100
20	1101	1200
21	1201	1300
22	1301	1400
23	1401	1500

24	1501	1750
25	1751	2000
26	2001	2500
27	2501	3000

Table 3. Flag values after the quality control

Flag	Comments
1	Correct value
2	Inconsistent value (out of narrow range limits)
3	Dubious value (spike)
4	Dubious value (vertical instability)
5	Not used
6	Inconsistent value (out of broad range seasonal limits)

Table 4. Projects within the IBAMar database

Major Project	Number of Cruises	Years
BALEARES	8	1985–1988
BALEARIC TUNA	14	2001–2012
CANALES BALEARES	13	1990–1993
CANALES-MATER	17	1996–1998
CIRBAL	30	1982–2006
COSTA BALEAR	39	1974–2011
COSTA CATALANA	2	1984–1984
EVALUACION ACUSTICA	5	2008–2012
IDEAS	9	2002–2010
RADMED	27	2007–2012

Table 5. Spatially averaged water properties in three zones of the Western Mediterranean

Water Mass	Temperature	Salinity	Oxygen	Nitrate	Nitrite	Phosphate	Silicate	Chlorophill-*a*
	(º C)	(PSU)	(ml l^{-1})	(µM)	(µM)	(µM)	(µM)	(mg m^{-3})
Alboran Sea								
Coast								
Surface water (< 20 m)	18.16±3.26 (264)	37.13±0.46 (264)	5.05±0.47 (195)	0.64±1.05 (76)	0.1±0.11 (76)	0.07±0.07 (76)	1.04±0.66 (76)	0.74±0.83 (78)
Open								
Surface water (< 20 m)	18.1±2.76 (308)	36.93±0.4 (308)	5.27±0.34 (220)	0.64±1.37 (109)	0.09±0.14 (109)	0.07±0.08 (109)	0.97±0.95 (107)	0.7±0.97 (107)
LIW (300 - 500 m)	13.24±0.04 (224)	38.5±0.03 (224)	3.81±0.2 (166)	9.34±2.67 (32)	0.05±0.04 (32)	0.47±0.17 (32)	7.68±1.98 (32)	0.36±0.16 (2)
South Balearic Islands								
Coast								
Surface water (< 20 m)	20.45±4.04 (920)	37.61±0.28 (914)	5.02±0.51 (513)	0.11±0.14 (146)	0.05±0.04 (145)	0.1±0.08 (145)	0.73±0.22 (145)	0.16±0.13 (103)
Open								
Surface water (< 20 m)	21.63±3.8 (1693)	37.54±0.32 (1692)	4.89±0.37 (1047)	0.17±0.56 (229)	0.05±0.07 (222)	0.11±0.08 (222)	0.76±0.39 (222)	0.21±0.22 (83)
LIW (300 - 500 m)	13.26±0.08 (1410)	38.5±0.03 (1396)	3.93±0.23 (908)	7.28±1.58 (77)	0.07±0.09 (77)	0.44±0.1 (75)	6.65±1.14 (77)	
WMDW (> 1500 m)	13.15±0.04 (47)	38.47±0.01 (47)	4.37±0.16 (43)	8.22±1.26 (6)	0.03±0.02 (6)	0.42±0.08 (6)	8.46±0.48 (6)	
North Balearic Islands								
Coast								
Surface water (< 20 m)	19.43±3.92 (1731)	37.7±0.29 (1626)	5.1±0.46 (867)	0.91±1.6 (408)	0.13±0.29 (401)	0.11±0.13 (402)	0.76±0.34 (304)	0.25±0.23 (200)
Open								
Surface water (< 20 m)	19.38±4.11 (2190)	37.8±0.31 (2188)	5.07±0.39 (1221)	0.2±0.49 (626)	0.07±0.14 (618)	0.09±0.12 (618)	0.86±0.46 (618)	0.31±0.45 (310)
LIW (300 - 500 m)	13.21±0.11 (1970)	38.49±0.05 (1955)	4.08±0.26 (1106)	7.9±1.56 (276)	0.08±0.12 (277)	0.41±0.13 (277)	6.49±1.17 (277)	
WMDW (> 1500 m)	13.17±0.04 (56)	38.48±0.01 (56)	4.41±0.17 (51)	7.8±1.65 (14)	0.08±0.14 (14)	0.41±0.07 (14)	7.96±0.86 (14)	

(Article history: Received 6 June 2014, Accepted 12 November 2014, Available online 6 January 2015)