myOcean

Project: WP15

Implementation of quality control procedures by means of probabilistic estimate of Mediterranean Sea temperature and its temporal evolution

First release:
2010

Authors for deliverable:
Marco Gambetta, INGV, Italy
Giuseppe M.R. Manzella, ENEA, Italy

Table of contents

Abstract

The quality assurance - quality control processes assure the quality of data. In particular quality control for ships of opportunity has been implemented in the Mediterranean during the Mediterranean Forecasting System - Pilot project. During the year a significant amount of data have been collected and the existing temperature data base in the Mediterranean is now capable to provide more information on temporal and spatial variability at monthly and mesoscale. During the last years the original Ships Of Opportunity Programme in the Mediterranean has been complemented with a so called ‘Cruises Of Opportunity Programme’, i.e. research vessels were used to collect data to be used in near real time. This has required the implementation of the original quality control procedures, since also CTD data have been transmitted in near real time. We started to modify all the procedures and constructed new temperature profiles to be used for comparison with new acquired data. Best estimates of monthly temperature profiles in 1x1 degree boxes have been calculated by using a maximum likelihood approach applied to data statistical distribution. The maps obtained from these estimates are compared with existing climatologies. Also the temporal variability of the temperature is calculated in different areas. This work is at the base of a further implementation of the quality control in the Mediterranean for both near real time and historical data. These are now also compared with the 'best estimates' and the consistency of the data will be assessed not only in the contest of each single campaign, but also on the base of the temporal variability in the areas. These information can be included in metafiles to be used for re-analysis and studies on climate variability. The improved procedures have a paramount importance for the management of data transmitted by observatories composed by different platforms.

1 Introduction

MyOcean is a three year project (2008 - 2011) having the goal to develop and validate the GMES Marine Core Service for ocean monitoring and forecasting. It is a transition project that will introduce more integration, operationality, and services in the existing European systems. In MyOcean, the R&D efforts will enable major progresses along the temporal dimension improving short-term predictions based on the best possible atmospheric forcing functions and re-analyzing the past ocean variability at temporal scales from seasonal to decadal. These activities are mainly carried out by means of models, however they necessitate of remotely sensed and in situ data.

Forecast, analysis and reanalysis need a base of in situ data of high quality. They are assimilated in numerical models in order to provide products on the past or future ocean characteristics. To have good data it is necessary to apply precise methodologies and protocols during all the phases of data collection from preparation of instruments (e.g. calibration) to post processing. Normally these activities are divided in two wide processes that take the names of Quality Assurance and Quality Control.

The best quality assurance practices have been specified in a Eurofleet report (Acquisition quality assurance procedures and data transmission) and the general model is presented in figure 1.

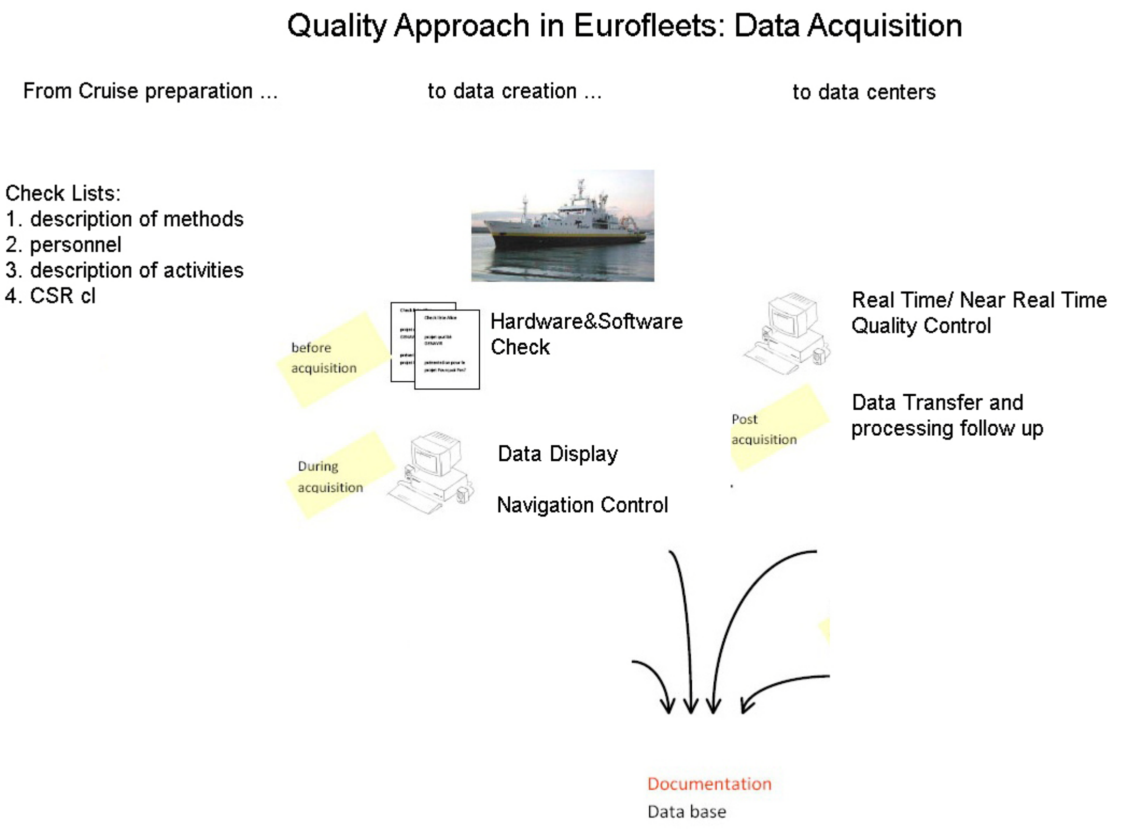


Figure 1. Best practices for quality assurance (From Eurofleet report WP10.7).

Quality assurance procedures include the preparation of personnel, test of instruments, calibration/intercomparison when possible, control of data and instruments during acquisition.

Post processing, is the process starting after data acquisition and is composed by steps described in many protocols and documents approved at international level. An important library is provided by IODE (http://www.oceanteacher.org/OTMediawiki/index.php/Online_Marine_Books).

It has been recently realised that the temperature data are affected by many errors. The causes are due to many causes: the technological changes and precision of instruments, the lack in some case of precise quality assurance - quality control procedures and information on how the data were obtained (metadata). In particular, there are uncontrolled errors that are associate to the use of XBTs, which represent the major percentage of temperature data in the Mediterranean as well as in the world Ocean (e.g. Gouretski and Koltermann, 2007).

The errors in data have affected their use for the estimation of climate variability/changes and climatologies (Levitus et al., 2000). For the Mediterranean the use of data affected by errors changing from year to year is even more problematic, due the large seasonal and interannual variability in the area. Normally the mean characteristics have been evaluated by averaging data in regions or regular boxes. Many efforts have been done to define areas were the physical profiles have similar characteristics. Manzella et al. (2003) defined these areas in terms of mean and standard deviations, Maillard et al. (2001) defined homogeneous areas in terms of similarity among profiles. However it is easy to demonstrate that the averaged data does not correspond to that having the most probable value on the base of statistical distribution. This is shown in figure 2, data from 1 x 1 degree box in the Tyrrhenian Sea is shown. The yellow line is indicating the mean value 14.31 °C (normally what is calculated in a climatology), but the highest amount of data are distributed around the value of 14.07°C.

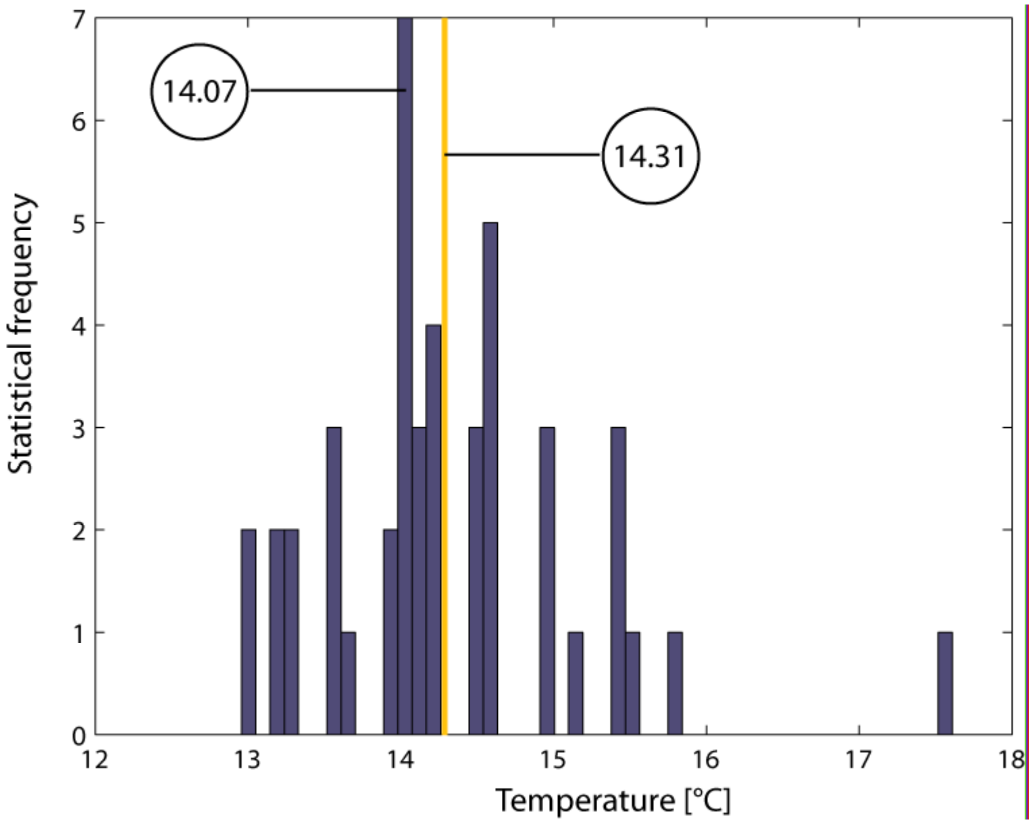


Figure 2. Statistical distribution of temperature values in bins of 0.5°C. The calculated mean value is indicated with a yellow line (see paragraph 2 for details).

It can be noted also that the distribution is not normal and will be shown that in some areas there is a multimodal distribution of the data. This is due to many causes, among which the uneven spatial and temporal distribution also in relatively small cells, the undersampling, and changes in circulation and physical characteristics of the sea. Our conclusion is that averaged data are then not providing the best estimate of the sea characteristics for the assessment of the quality of the data, that could be better done by estimating profiles on the base of their statistical distribution.

The implementation of new procedures of quality control for the near real time data has been required by the evolution of the Ship Of Opportunity Programme in the Mediterranean that was complemented with a so called Cruise Of Opportunity Programme, i.e. the use of research vessels to collect both XBT and CTD data to be assimilated in forecasting systems. The implementation has been done having also in mind that new multiplatform observing systems are under development (the 'observatories'). The idea is to have procedures to be applied in near real time to many different parameters, in an identical way of delayed mode procedures. An additional qualitative test has been added looking at the temporal variability of temperature data, and comparing this to the data to be processed. In this way we will be able to provide additional information on data quality, specifying the 'similarity' with best estimates, what is the qualitative behaviour of data with respect to long term variability, and what is the quality on the base of precision of instrument used for acquisition.

2 Material and method

2.1 Material

Data used for the construction of the temperature climatology are principally derived from MEDAR / MEDATLAS, MATER and MFS. Additional data from MedGOOS have been added. The data have been collected during 12912 cruises, and provide 24229158 unique sea water temperature values from September 1900 to October 2009 (Figure 3). The total amount of available casts has a temporal distribution in the years and months as shown in figure 4 and table 1.

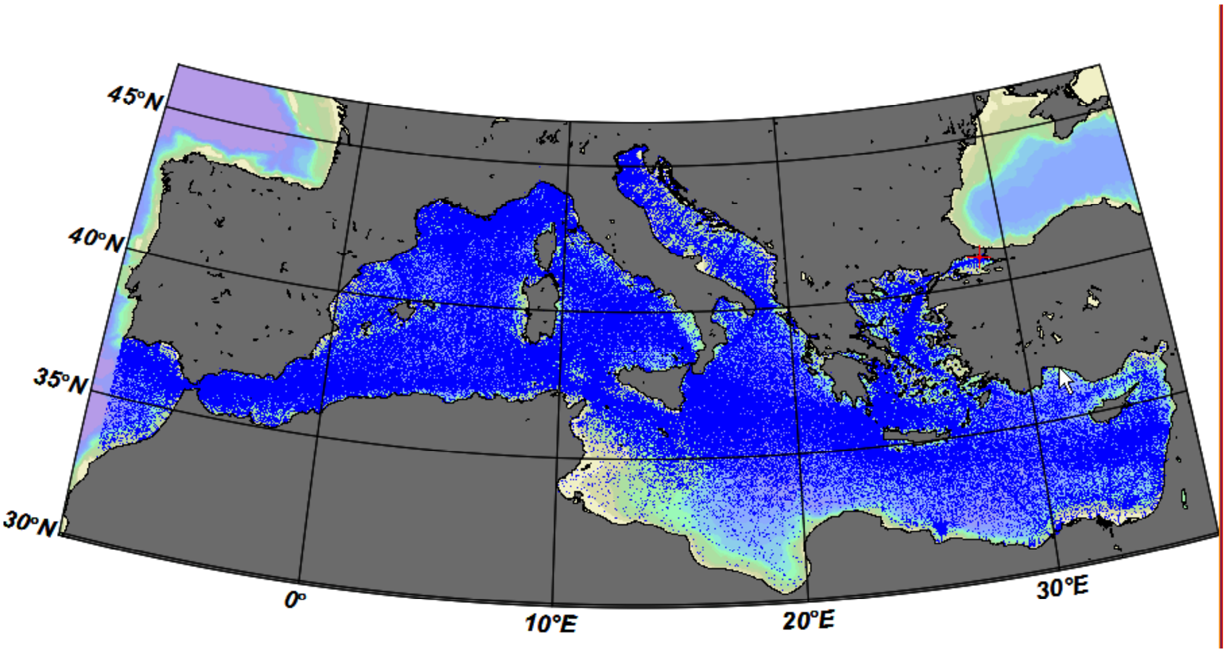


Figure 3: Spatial distribution of historical data available for calculation. A blue dot represents a single cast location

The data have been collected with different methodologies and technologies in about one century. The initial temperature and salinity data were collected in the Mediterranean at the beginning of XX century. From an historical point of view, it is important to remember the Danish expeditions Thor (1908-1910) and Dana, this one being part of the world wide 'Carlsberg Foundation's Oceanographical Expedition round the World 1928-30' (the official name of the expedition). During a detour into the Mediterranean the equipments on board the Dana were tried out and data were collected in eighteen stations. Other important initiatives were promoted by the Trieste Thalassographic Experimental Institute, an Austro-Hungarian scientific institution founded in 1841, that was assigned to the Italian Thalassographic Committee (founded in 1909) after the first world war.

A long period without expeditions in the Mediterranean lasted for about five years during the 1914 - 1918 world war. The data collected in the period 1921 – 1924 were essentially part of exploration conducted locally in the Alboran Sea and the area around the Strait of Messina. A significant amount of data was collected in the Mediterranean starting from 1943, in an area between Gibraltar and Sicily. These data collection was probably related to the preparation of invasion of Sicily by the allied navies. In 1944 the data were collected in all the western Mediterranean, probably for a similar war reasons. It is important to remember also some important programs that allowed to know the characteristics of the water masses: the International Geophysical Year (1957 – 1958) established in the framework of ICSU (International Council for Science), Gibraltar and MEDOC Experiments (1968 – 1972; e.g. Lacombe et al, 1985; MEDOC Group, 1970), established in the framework of CIESM (International Commission for the Exploration of the Mediterranean Sea), Western Mediterranean Circulation Experiment (1985 – 1986; WMCE, 1989) and Physical Oceanography of the Eastern Mediterranean (1986 – 1987; e.g. Robinson et al., 1991), the last one conducted in the framework of IOC programmes.

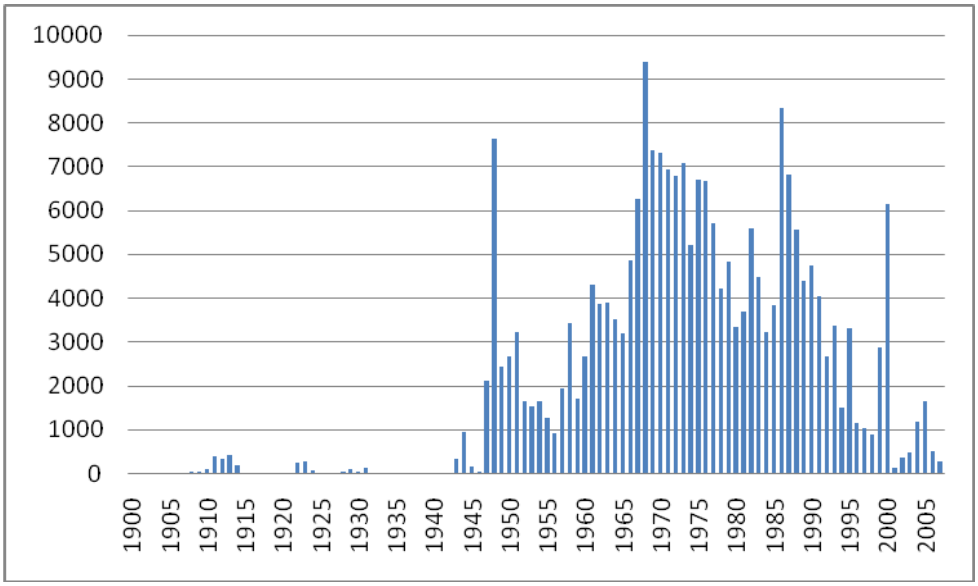


Figure 4. Temporal distribution of data during the years

The number of data used for the construction of the climatology is summarized in the table 1.

Month	Initial Year	Last Year	Number of cruises	Number of casts	Number of data points
January	1908	2007	472	14613	1909871
February	1908	2007	615	19598	1933430
March	1908	2006	618	22523	2402498
April	1908	2007	612	21016	2026822
May	1908	2007	673	29200	2619320
June	1908	2007	656	23299	1854851
July	1901	2006	606	22560	1203863
August	1910	2009	580	21516	1354351
September	1900	2009	659	22146	2131501
October	1914	2009	602	24502	2973242
November	1911	2006	577	21659	2708657
December	1908	2006	462	12912	1523574
Total			12912	251706	24229158

Table 1: Temporal data distribution, data of cruises headers

From this short presentation, it appear that the historical data are scattered in space and time, and have been collected with very different instruments such as bottles, STDs, CTDs, MBTs, XBTs, each of them having a different precision, as summarised in table 2 for the temperature sensor. Today is not easy to know the precision to be assigned to data from each single cruise, since many information (now called metadata) have been lost. In a general way, it can be assigned an approximate precision on the base of the years during which the data were collected. This is a working hypothesis, shown in table 3 for temperature and salinity, that has been used in order to roughly define the quality of the data.

Instrument	Temperature precision	Depth precision	Year (from)
Nansen Bottles	0.01 °C	1.5-6% FS	1897
STD (Howe and Tait, 1965)	0.1 °C	?	1965
CTD (Neil Brown	0.001 °C	0.015% FS	1967
MBT	0.2 °C	1% Z	1940
XBT (Robinson)	0.1 °C	2% Z	1966

Table 2. Precision of instruments normally used for temperature collection. STD stand for Salinity, Temperature, Depth probe, CTD for Conductivity, Temperature, Depth probe, MBT for Mechanical Bathythermograph, XBT for eXpandable BathyThermograph. Temperature units are in degrees Celsius, Z stands for depth and FS for Full Scale. Basic references on instruments are provided in parenthesis.

Parameter	1908-1970	1970-1980	1980-2009
Temperature	0.01	0.05	0.005
Salinity	0.02	0.1	0.002
Density	0.02	0.1	0.002
Depth	5 m	1 dbar	0.5 dbar

Table 3. Estimated precision during different decades(qualitative estimate).

With respect to previous studies (e.g. MEDAR Group, 2002; Rixen et al., 2005; Manzella et al., 1990), the data base used for this study is containing more data derived from MFS-VOS projects that cover in particular especially the eastern and southern Mediterranean from 1999 (Pinardi and Coppini, 2007). The working hypothesis of table 3 does not hold for MBT and XBT data, the associated error being revised in new studies (e.g. Reseghetti.,), as part of the Mediterranean Operational Oceanography Network activities.

2.2 Method

In order to achieve a reliable set of climatological sea water temperature profiles the Mediterranean sea has been subdivided in monthly 1x1 degree spatial boxes. The historical data have, among others, a different vertical resolution. Automated probes normally have a software providing data at (say) 1m interval, bottles data were collected a pre-defined depths depending on the particular research purposes. Till the 70s, the MBT and XBT data were derived from graphs and the quality was greatly depending on the ability of the technician to transform the graphs in digital records. In some cases the original data have been lost and only data interpolated at standard levels are available. This situation makes difficult a simple calculation of mean profiles. Figure 5A provide an example of profile existing in one particular area (namely the Tyrrhenian Sea). The simple average of all data would produce the profile in figure 5B. The standard deviation associated is very large and noisy (not shown).

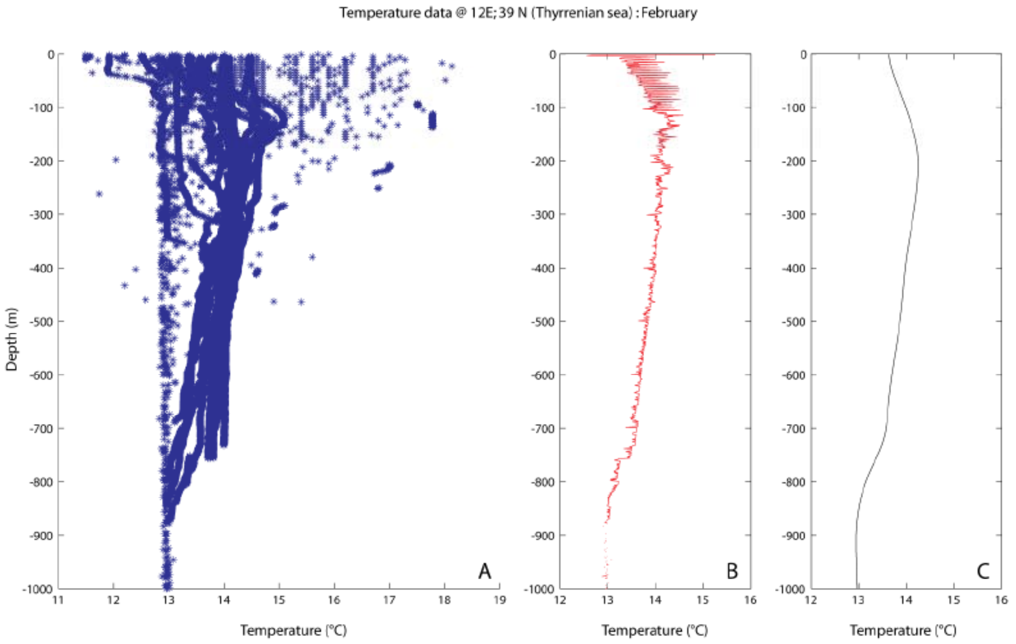


Figure 5. The original data in a particular 1x1 degree cell (A), the mean profile obtained by simply data average (B) and a possible expected man profile (C)

Looking carefully at all the profiles for a particular month, it is possible to see different possible 'mean' profiles. This could indicate the presence of different flow regimes, varying spatially and temporarily from year to year or the presence of wrong data. Assuming that different flow regimes are present in the data, two kind of mean profiles can be defined, a first more vertically homogeneous centred on the temperature value of 13°C, a second similar that shown in figure 5C. It will be shown that in practice, more than one mean profile could be defined in some cells, each of them being representative of different circulation patterns.

Due to the presence of profiles whose quality is questionable, the 1x1 degree boxes data have been checked for quality and consistency, applying a threshold rejection algorithm: data out of selected ranges (5° - 35°C at basin level) were deleted. The number of deleted profiles was about 1% of the entire original data base. Within each box a statistical distribution for bins of $\delta T = 0.025°C$ and $\Delta Z = 2$ m was calculated

For each grid node is defined a temperature $T_{i,j}$ and a range δ

$$P_{i,j} = \frac{\sum_{k=1}^{n} t_k}{m} \qquad (1$$

$$T_{i,j} - \delta \le t_k \le T_{i,j} + \delta \qquad (2$$

where n is the total amount of temperature values (observations) collected at the jth depth level and m is the number of tk that satisfy eq. 2 . The 2D Probability Space is then row-wise sampled to calculate the temperature from maximum distribution value. In vertical it is obtained a profile of temperature T_i that is not yet the best estimate. This is obtained from a parametric model $f(T_i|\theta)$, where θ the estimator. The maximum likelihood method provides many estimates of θ, the best one being obtained by maximizing or minimizing the likelihood. In this case the θ minimizing the RMS and vertical gradient was selected. The procedure is synthesized in figure 6, where the original data, the set of estimates and the maximum likelihood estimate are shown.

The best estimated profiles are smoothed and interpolated at 1 m vertical interval and standard deviations calculated. Each maximum likelihood profile is spatially associated to the geographical centroid of the observed dataset.

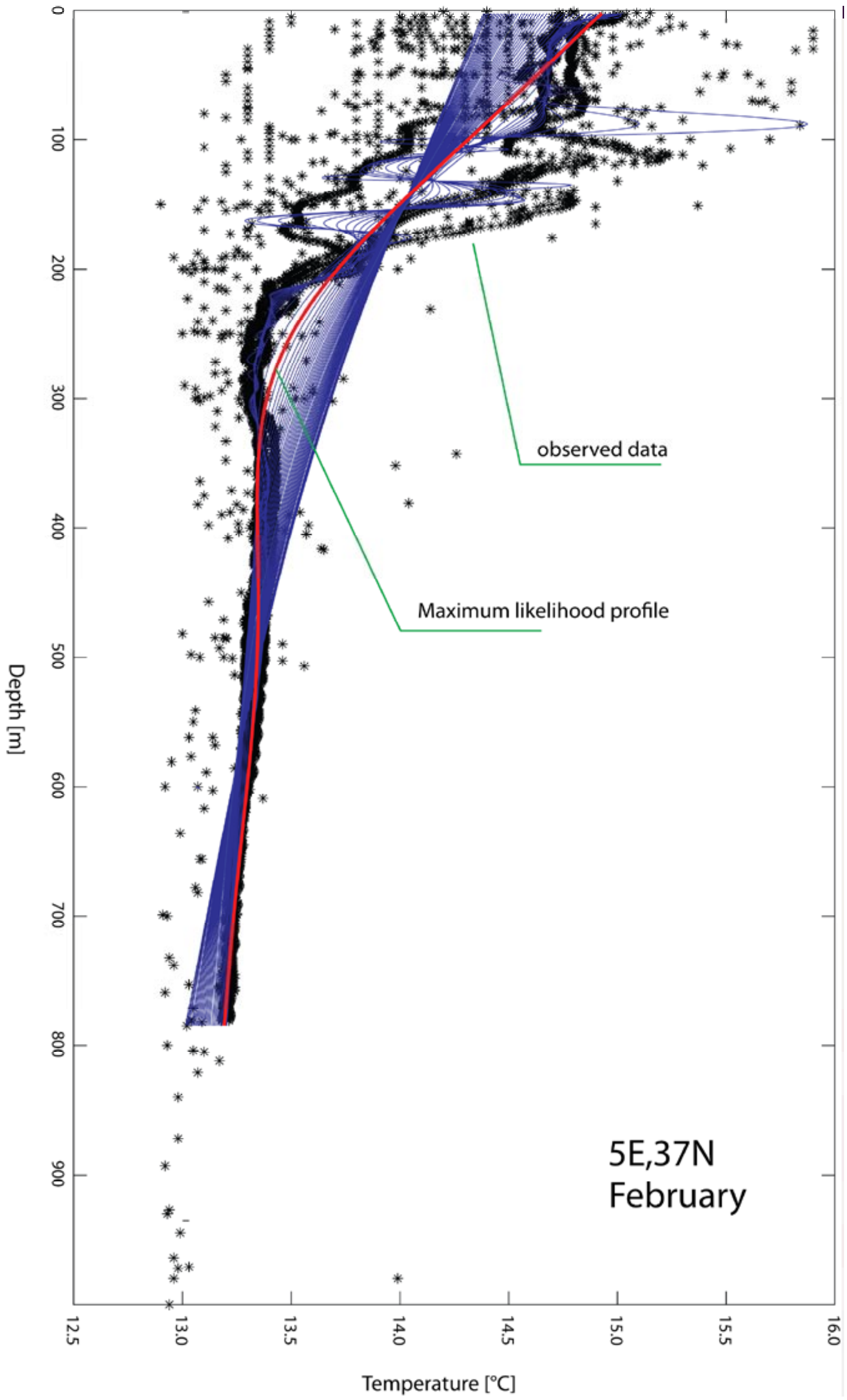


Figure 6. The maximum likelihood procedure for the best estimate of the temperature profile.

The monthly standard deviations are derived from the estimated temperature values and the original profiles. Since these ones were not having a regular vertical sampling, the standard deviation could have some values lacking at certain depths in some profiles; at depths where data were lacking the standard deviation was calculated by interpolation. A sampling indicator is also provided at each depth, in order to inform on data coverage within each cell (paragraph 2.4).

2.3 Other issues on statistical distribution

The above method implies that there is a unique main maximum in the data frequency distribution. However there are cases in which there are two maxima, and this obliges to introduce another assumption for the selection of the best estimate. Figure 7 provide an example of statistical distribution of data with two maxima

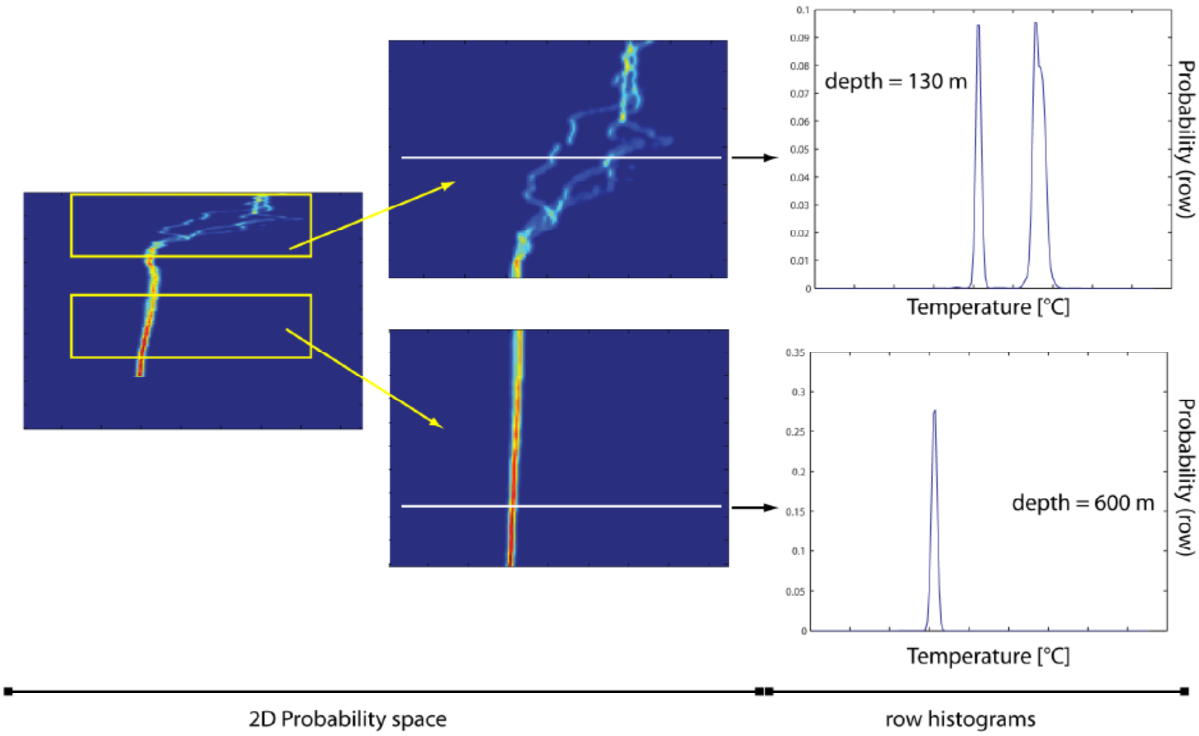


Figure 7: Sampling the 2D probability space

It may happen that a 2D Probability Space shows a multi modal distribution of the data. In this case a specialized algorithm finds and separates the maximum likelihood profiles accordingly to the data density distribution. The separation is performed analyzing the row-wise (i.e. : depth level) probability distribution.

2.4 Evaluation of the spatial sampling

An index of the dispersion of data points in each cell can be provided by calculating an average distance from each feature to its nearest neighbouring feature. Figure 8 is providing an indication on how data can be dispersed in a cell or potentially aggregated in clusters.

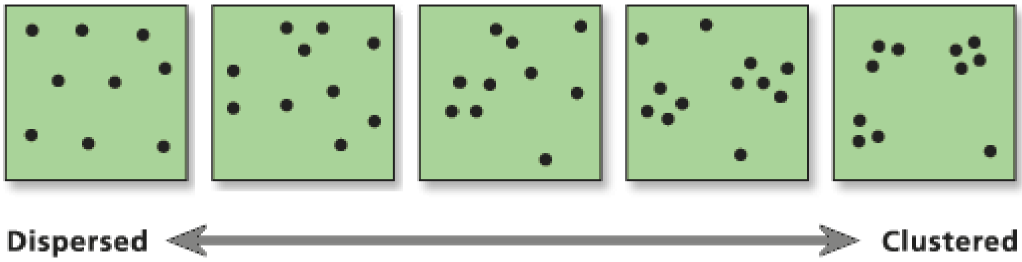


Figure 8. distribution of data points in a cell, from regularly dispersed to clustered

For definition, the nearest neighbour index is expressed as the ratio of the observed distance divided by the expected distance. The expected distance is the average distance between neighbours in a hypothetical random distribution. If the index is less than 1, the pattern exhibits clustering, and in particular zero is indicating that all data are represented by a single point; an index equal to 1 is indicating a random distribution; if the index is greater than 1, the trend is toward perfectly uniformly dispersed points. The maximum value can be 2.15, indicating the uniform dispersion.

A spatial sampling indicator is also provided as follows. First the mean next-neighbour distance $\bar{d}$ is calculated, then,s ince the box base is rectangular, the quantity $\bar{\delta}$ is calculated using eq. 3 (Davis 1973)

$$\bar{\delta} = \frac{1}{2}\sqrt{\frac{A}{n} + \left(0.514 + \frac{0.412}{\sqrt{n}}\right)\frac{p}{n}} \qquad (3$$

where A is the area of the cell, p is the perimeter, and n is the number of points within the cell. The expected ($\bar{\delta}$) and observed $\bar{d}$ mean next neighbor distances are used to construct an index (R) of the spatial pattern

$$R = \frac{\bar{d}}{\bar{\delta}} \qquad (4$$

This indicator ranges from 0:0 for distribution where all points coincide, to 1:0 for a random distribution of points, to a maximum value of 2:15. The latter value characterizes a distribution in which the mean distance to the nearest neighbor is maximized. This index is provided for each cell, depth and month to give an indication of the representativeness of the estimated temperature profile.

3 Mean monthly temperature fields

The Mediterranean Sea was covered with cells of 1x1 degree in latitude and longitude, and for each cell monthly profiles, standard deviation and data distribution indices were calculated. In some areas, where two distinct data convolutions were found, secondary mean profiles and standard deviations were calculated. Here the results for February and August are shown and discussed for the mean monthly profiles of the most probable temperature estimate, making comparisons with previous climatologies.

The temperature estimate at surface during February is shown in Figure 9a. There are some clear indications of well known circulation features such as the cold water in the Gulf of Lion with a pattern of isolines extending toward the Bonifacio Straits, relatively warm water in the easternmost part of the Mediterranean. The estimates can be considered only in deep sea and the cold water in the Adriatic Sea as well as relatively warm water in front of Venire are not considered representative of phenomena in the area.

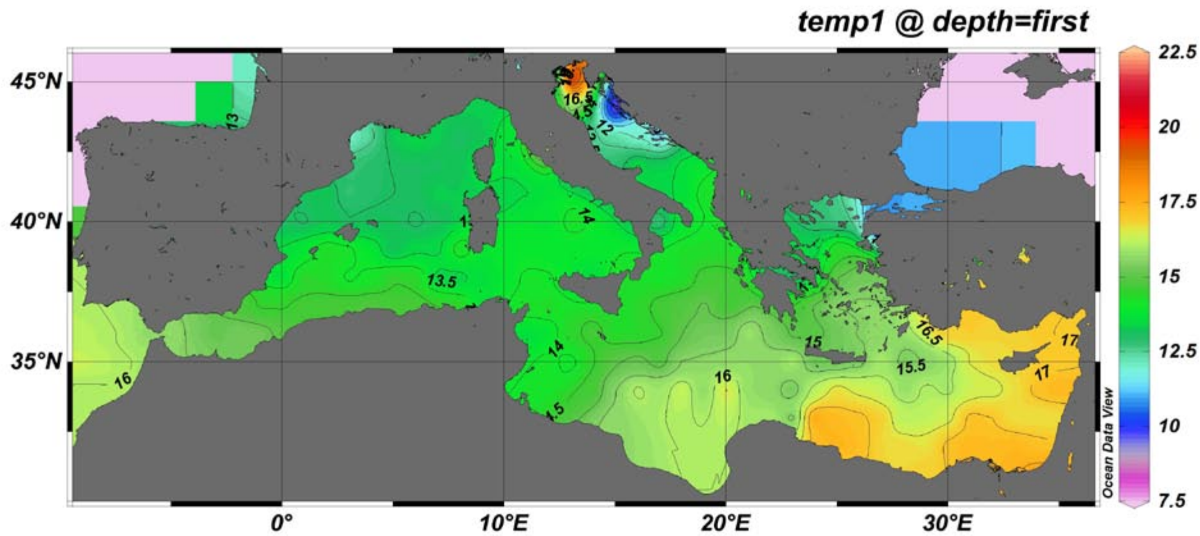


Figure 9a. Best estimate of surface temperature in February obtained for MyOcean (MedMyO-Aug). The temperature contour interval is 0.5 °C.

The comparison with the original Medar/MedAtlas climatology (figure 9b) is relatively useful, since this one is providing only a smoothed temperature field, without details on variability at spatial scale of 50 - 100 km.

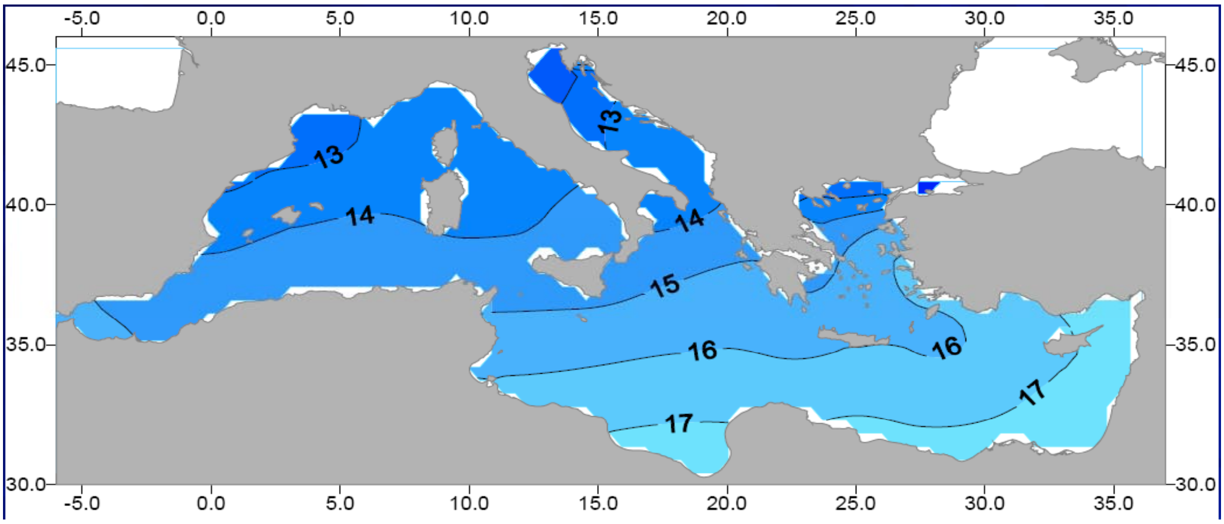


Figure 9b. Climatology of surface temperature in February from Medar/MedAtlas. Temperature contour interval is 1 °C.

A second comparison has been done with the Levitus 2009 climatology, that is presented in figure 9c. In this case the temperature field presenting more details than Medar/MedAtlas, but is still very smooth, and does not show the important activity in the Mediterranean at scales shown in the MyOcean climatology.

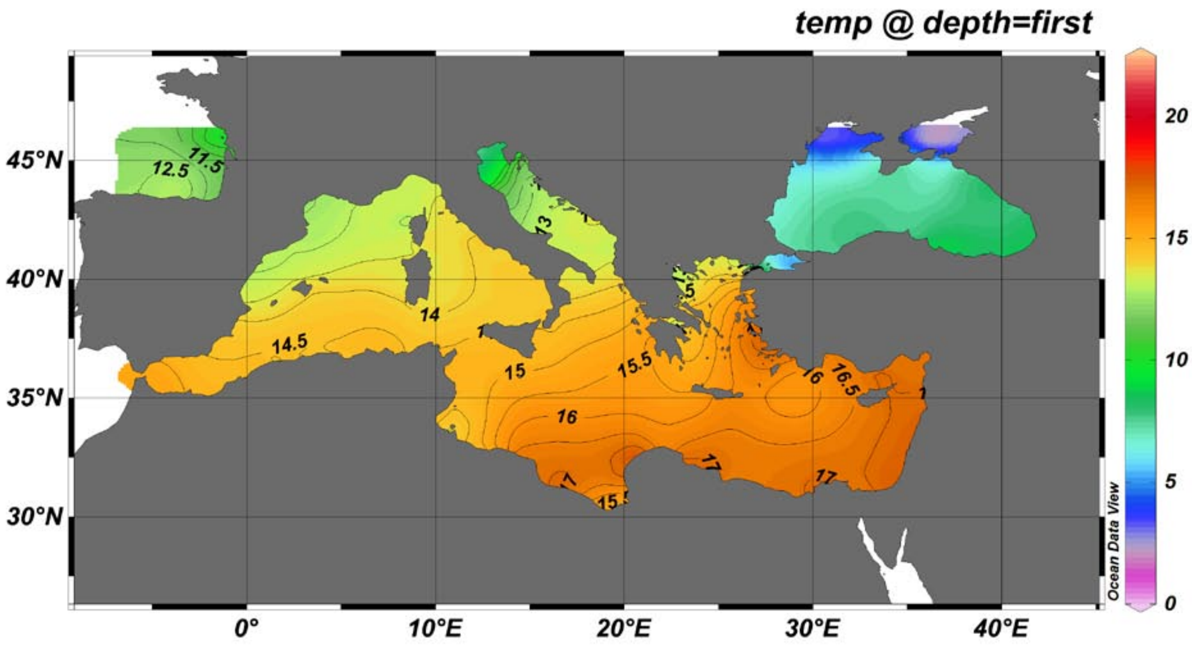


Figure 9c. Climatology of surface temperature in February from Levitus 2009. The temperature contour interval is 0.5°C.

Finally, a comparison was done with a recent climatology done in the framework of SeaDataNet project. The comparison is more interesting, since there are some common features and differences. The cold water in the Gulf of Lion and the presumable cyclonic circulation south of Rodhes, have similar patterns in both climatologies. The major differences are in the Adriatic Sea, probably due to the major availability of data in this region by Tonani. In the remaining part of the area the 'Tonani climatology' is smoother than the MyOcean.

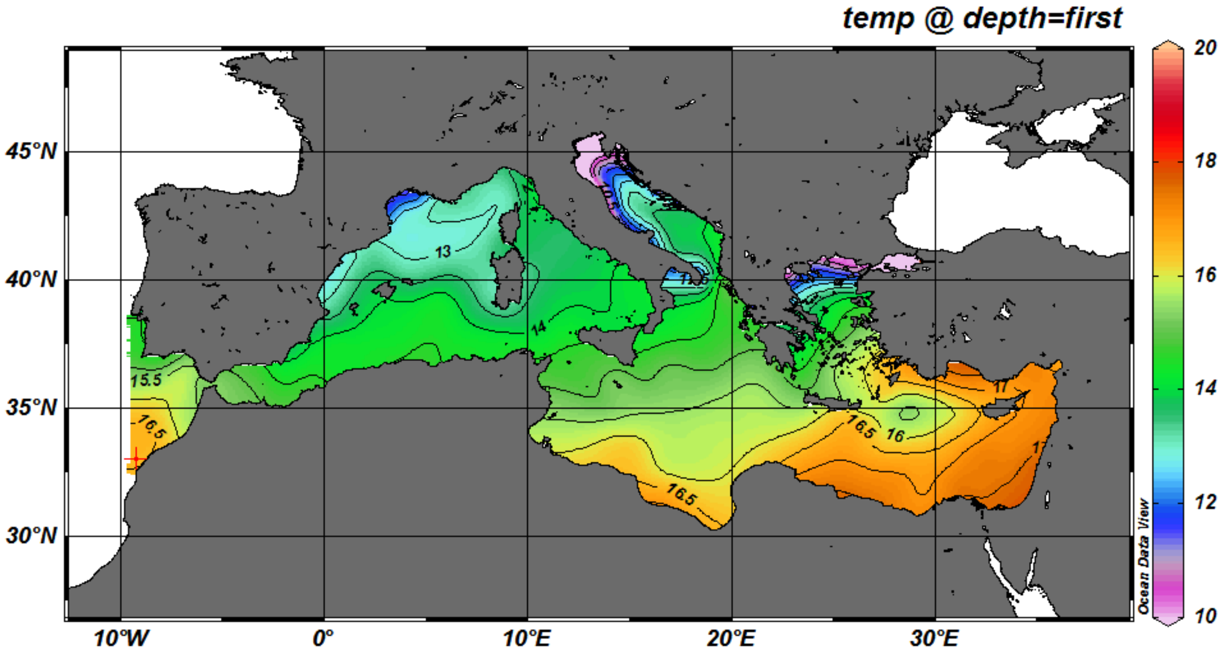


Figure 9d. Climatology of surface temperature in February from Tonani. Temperature contour interval is 0.5°C

In August the spatial gradients of the surface temperatures are higher than in February, as expected, and there are some typical patterns of the Mediterranean circulation, such as the relatively cold water in the Gulf of Lion, warm water in the Tyrrhenian Sea (figure 10a). Also in August the Medar/MedAtlas climatology is very smooth (figure 10b).

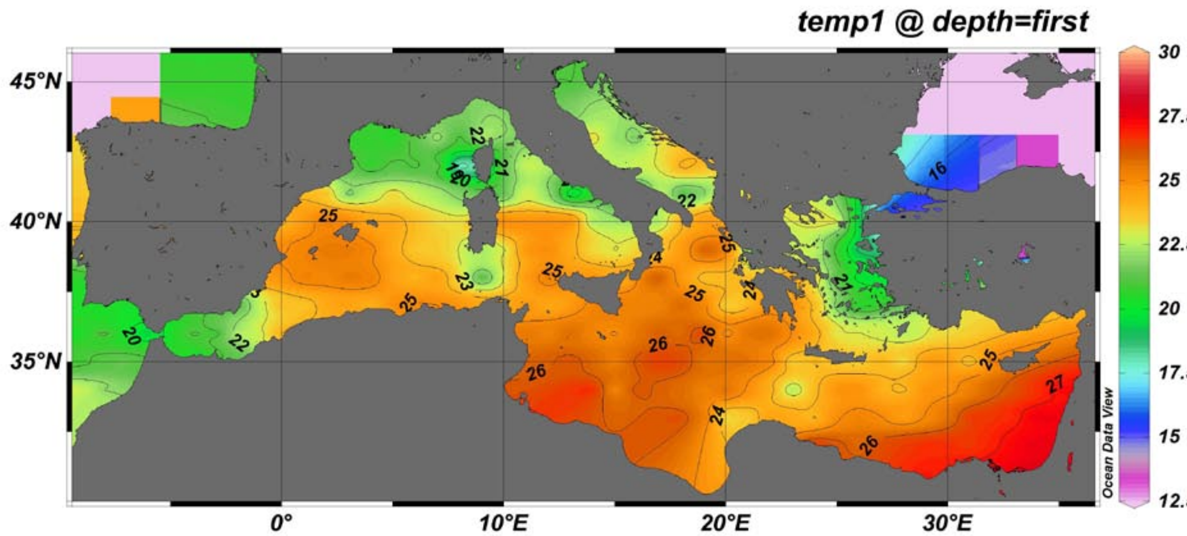


Figure 10a. Climatology of surface temperature in August obtained for MyOcean (MedMyO-Feb)

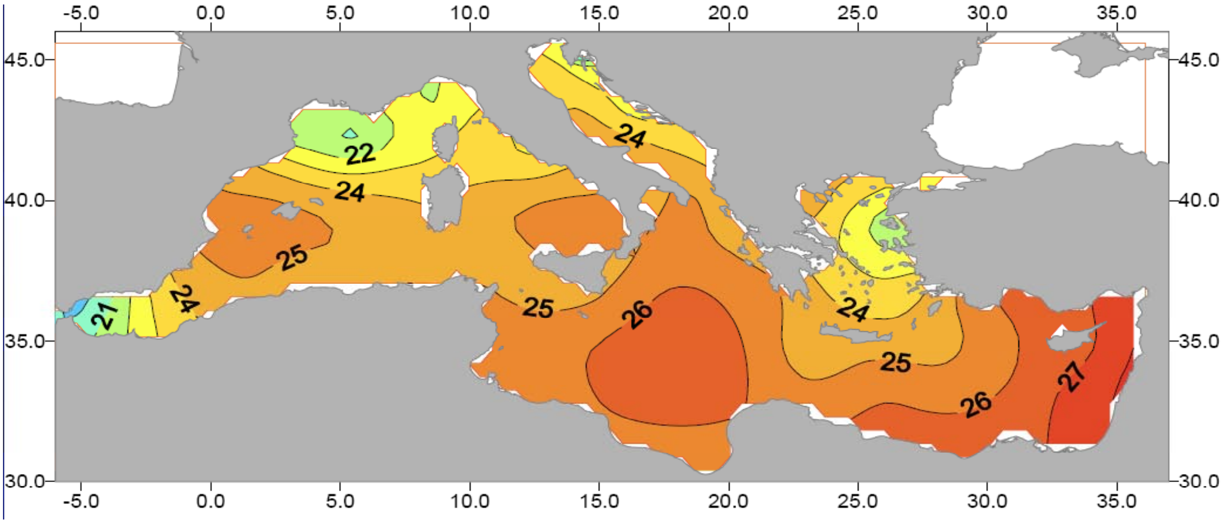


Figure 10b. Climatology of surface temperature in August from Medar/MedAtlas.

The Levitus 2009 climatology in February is presenting a higher spatial variability with respect to the MEDAR/MEDATLAS, but is in general quite consistent with the general patterns of the Medar/MedAtlas climatology.

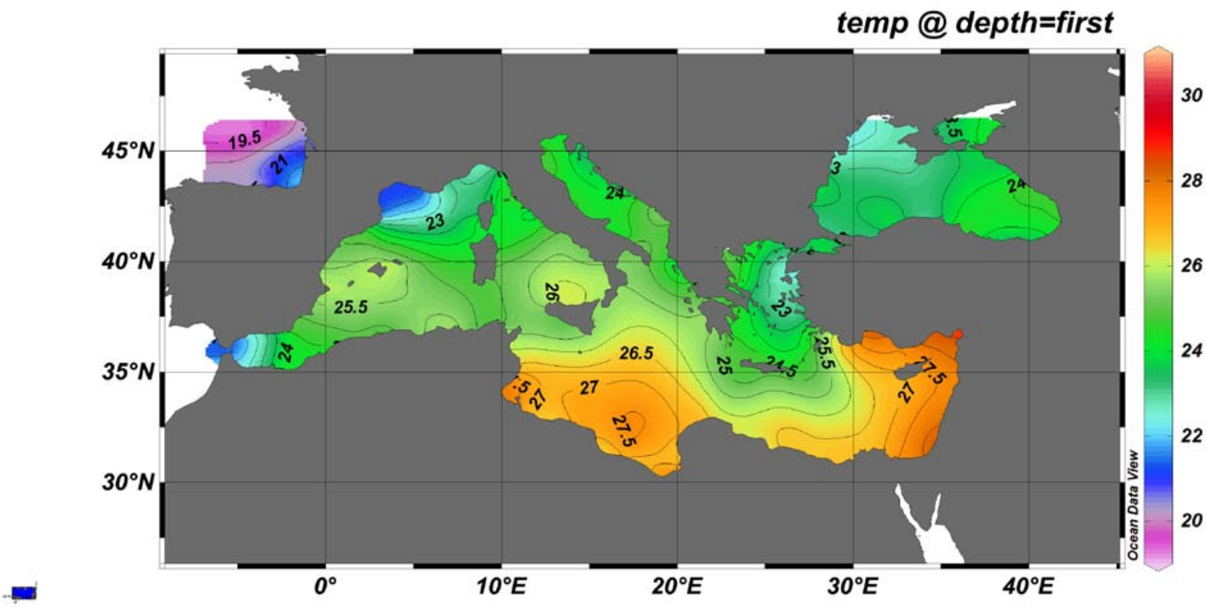


Figure 10c. Climatology of surface temperature in August from Levitus 2009.

The climatology by Tonani is having the same pattern of the Levitus 2009 climatology, but there are major differences in the Balearic area. Conversely, in this area there is a major coherence between Levitus 2009 and MyOcean.

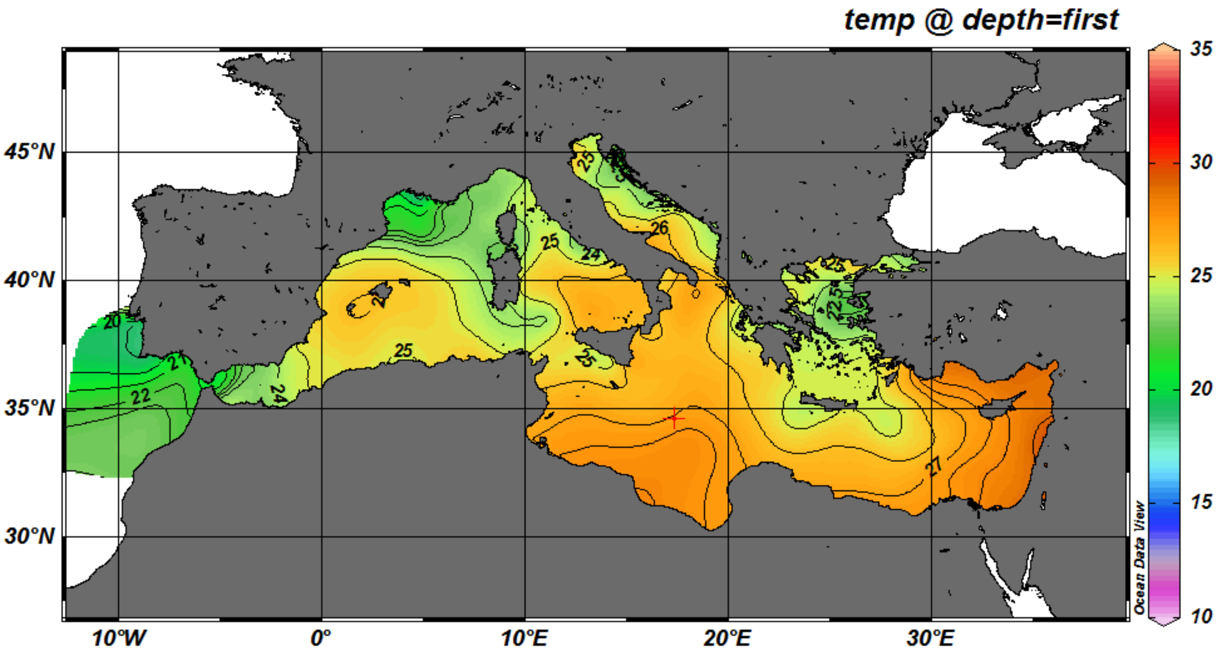


Figure 10d. Climatology of surface temperature in August from Tonani.

The monthly maps of surface temperatures are given in appendix 1.

4 Vertical behaviour of estimates

In the previous paragraph the main characteristics of the surface temperatures obtained from the best estimate have been described for selected months.

For particular areas, there is a bi-modal distributions of the data, indicating that another dynamical phenomenon is of importance for such areas. Figure 11 show the profiles of the two principal estimates for the month of February. It must be underlined that only about 20 over 256 boxes have significant bi-modal distributions.

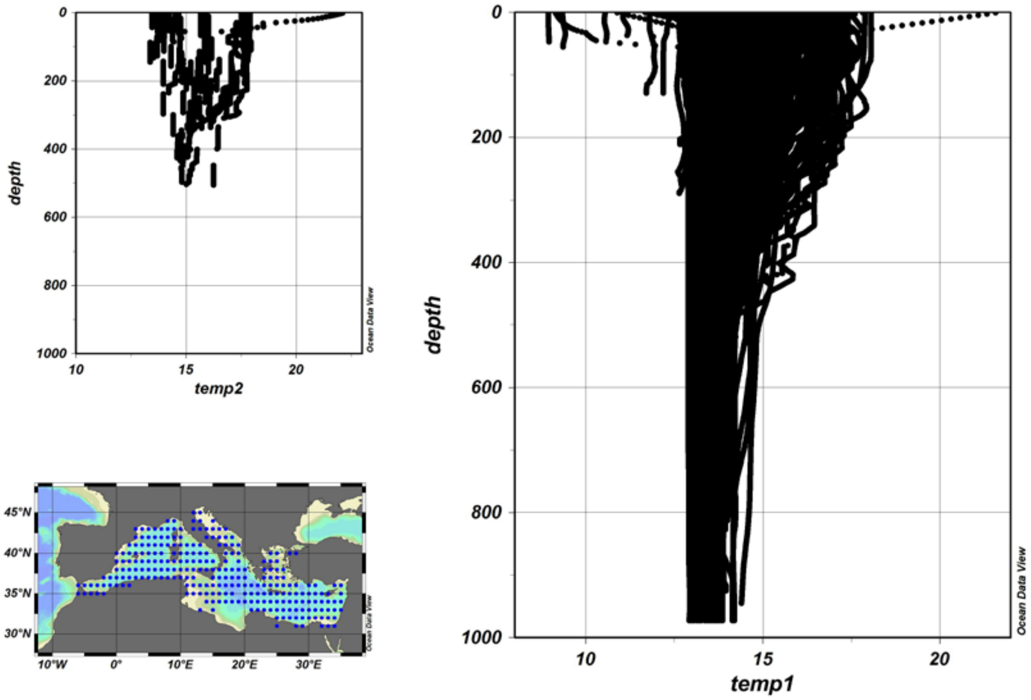


Figure 11. The best profiles estimated from statistical data distribution during February. Temp1 is the estimate with the major probability, temp2 was obtained in those areas were a bi-modal distribution was found.

In August about 30 over 264 boxes were presenting a bi-modal behaviour. The profiles are shown in figure 12.

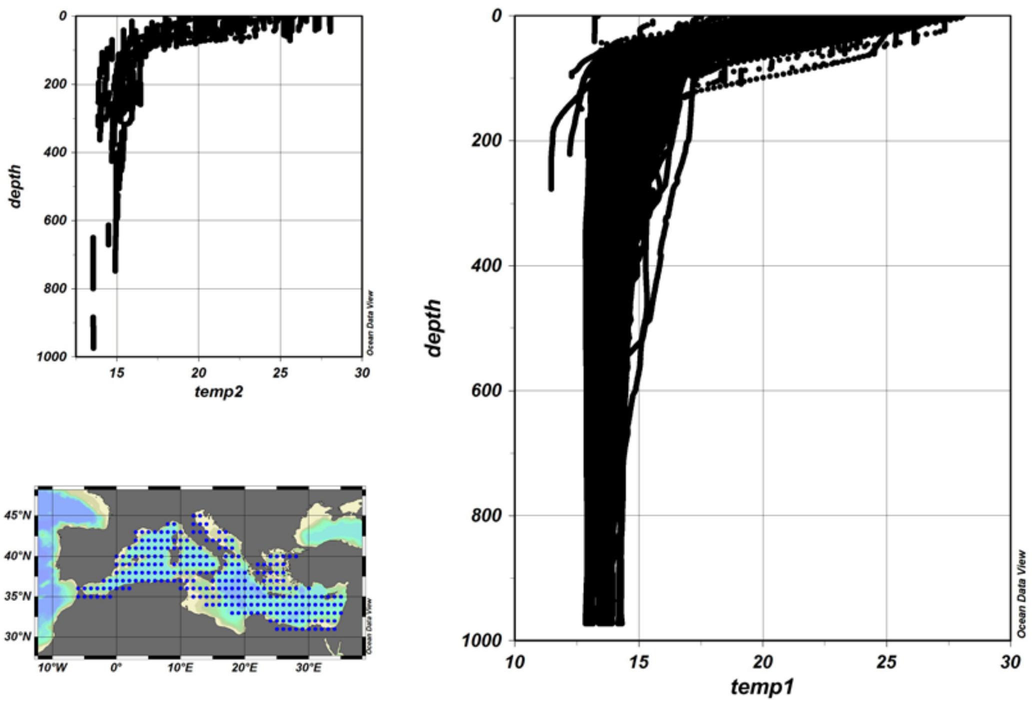


Figure 12. The best profiles estimated from statistical data distribution during August. Temp1 is the estimate with the major probability, temp2 was obtained in those areas were a bi-modal distribution was found.

5 Temporal variability

The temporal variability of the temperature in the Mediterranean has estimated by computing the mean monthly values for each year, when the number of profiles were significantly covering the basins considered. For the entire Mediterranean the variability during February at three different depth (surface, 300 m, 800 m) is shown in figure 13a.

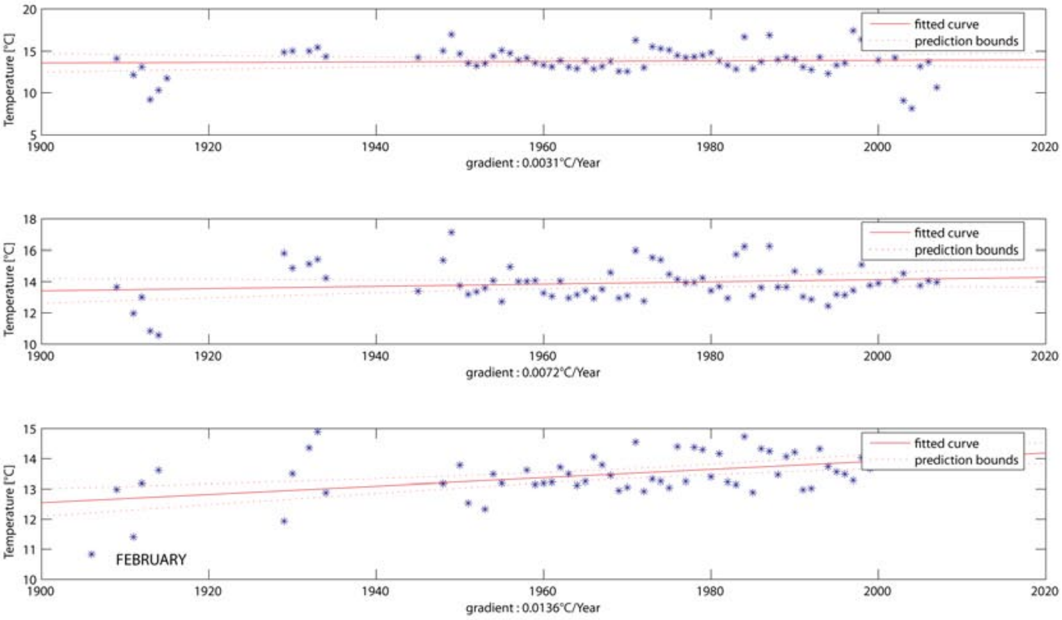


Figure 13a. The temporal variability of the temperature at surface, 300 m and 800 m during the month of February in the entire Mediterranean.

The same was done for the month of August and the result is shown in Figure 13b.

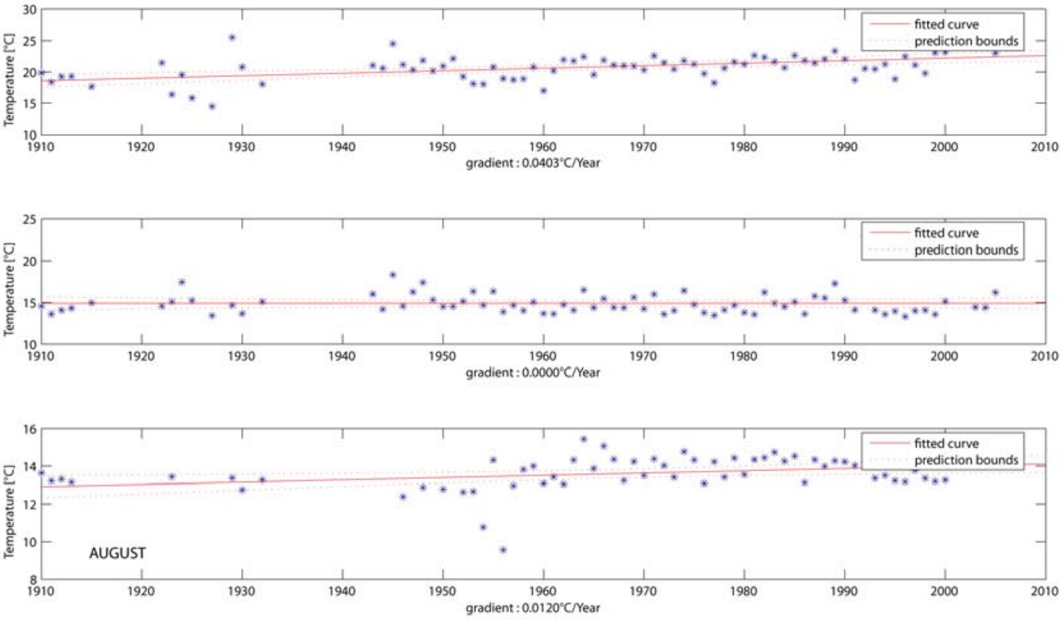


Figure 13b. The temporal variability of the temperature at surface, 300 m and 800 m during the month of August in the entire Mediterranean.

In both months there was a consistent increase of temperature at surface and in the deep waters, but not at intermediate depth.

The same analysis has been done for different basins, as shown in the figures below.

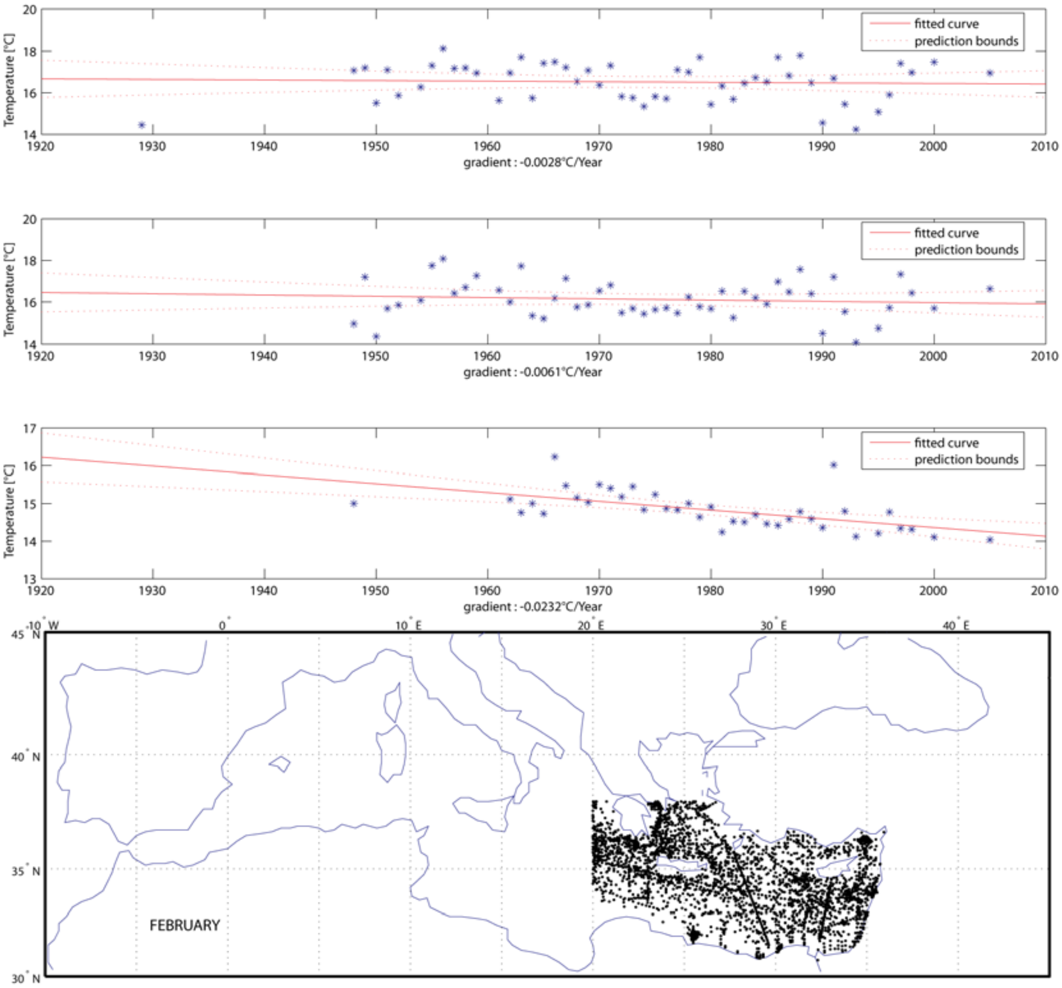


Figure 14a. The temporal variability of the temperature at surface, 300 m and 800 m during the month of February in the Eastern Mediterranean.

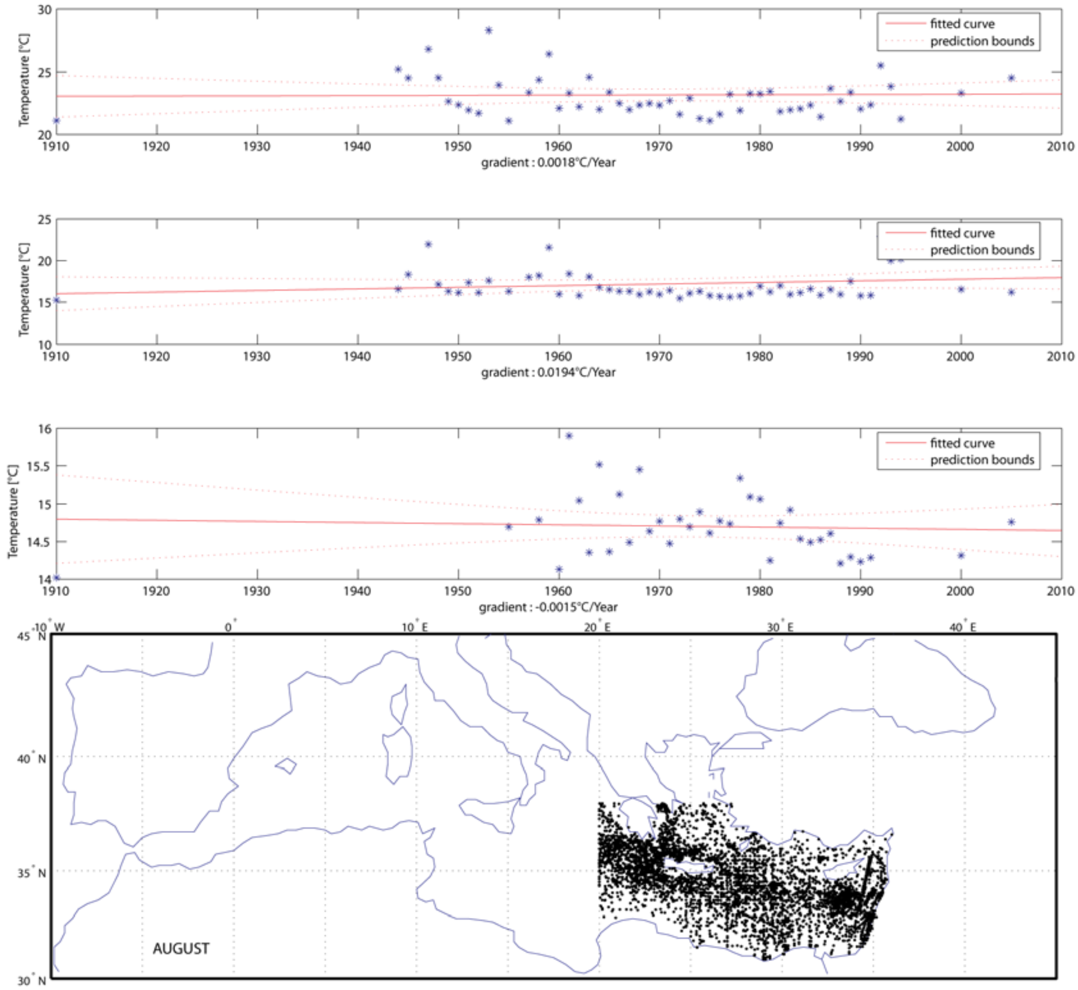


Figure 14b. The temporal variability of the temperature at surface, 300 m and 800 m during the month of August in the Eastern Mediterranean.

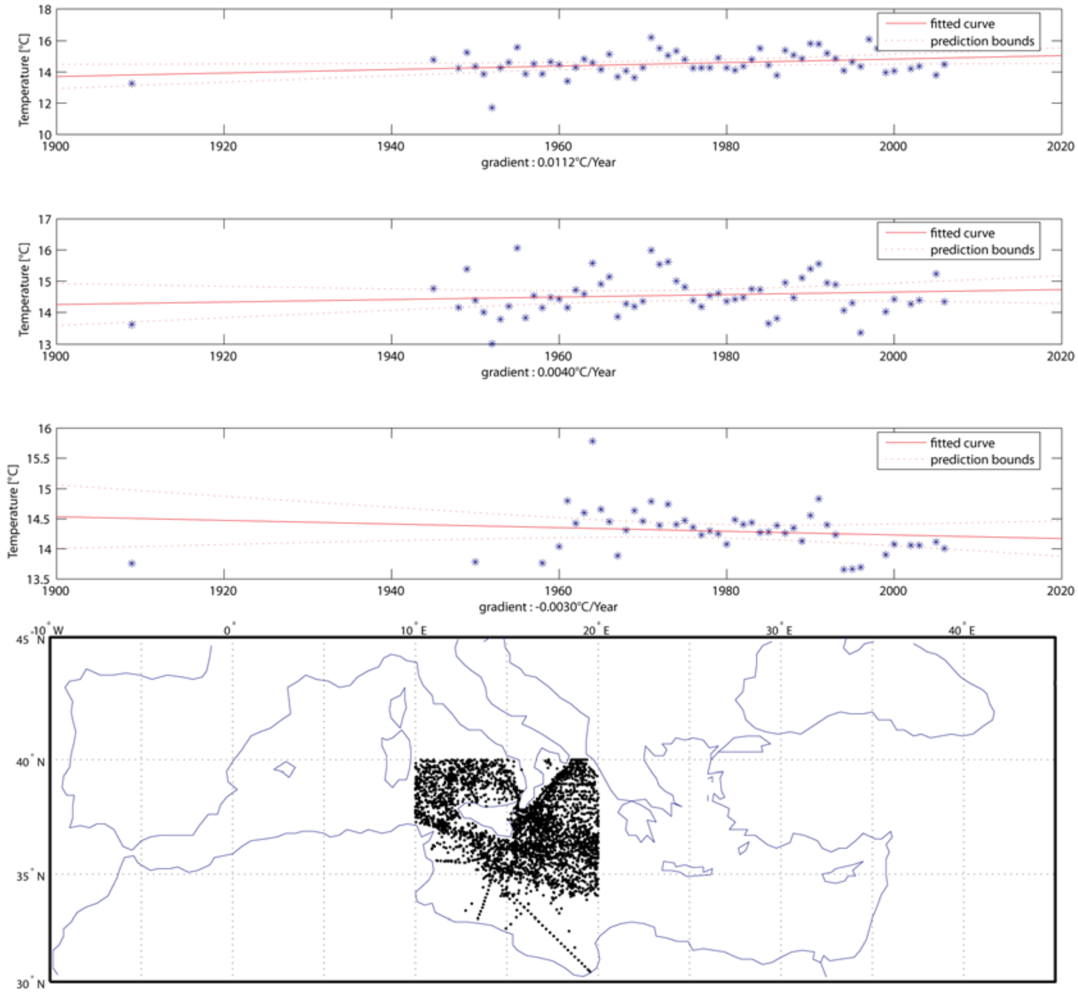


Figure 15a. The temporal variability of the temperature at surface, 300 m and 800 m during the month of February in the central Mediterranean.

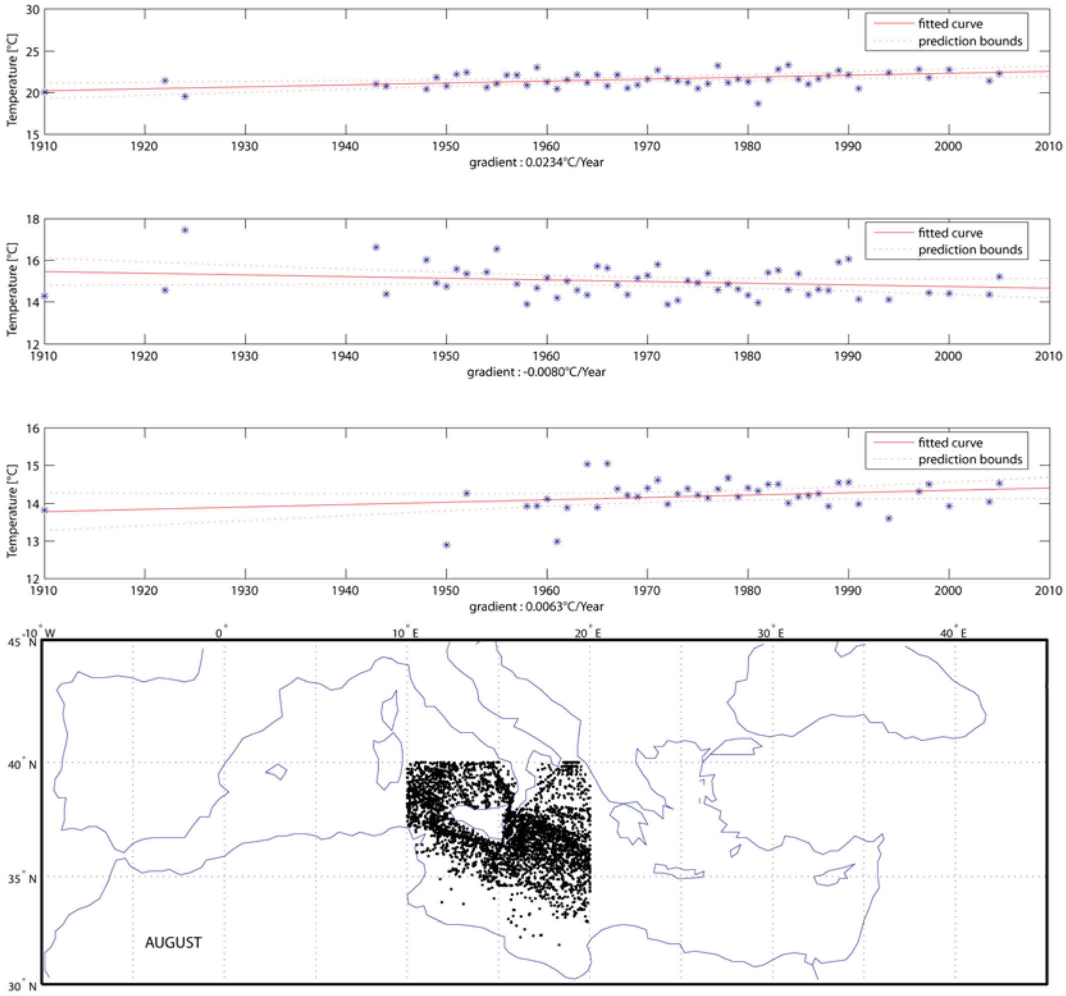


Figure 15b. The temporal variability of the temperature at surface, 300 m and 800 m during the month of August in the central Mediterranean.

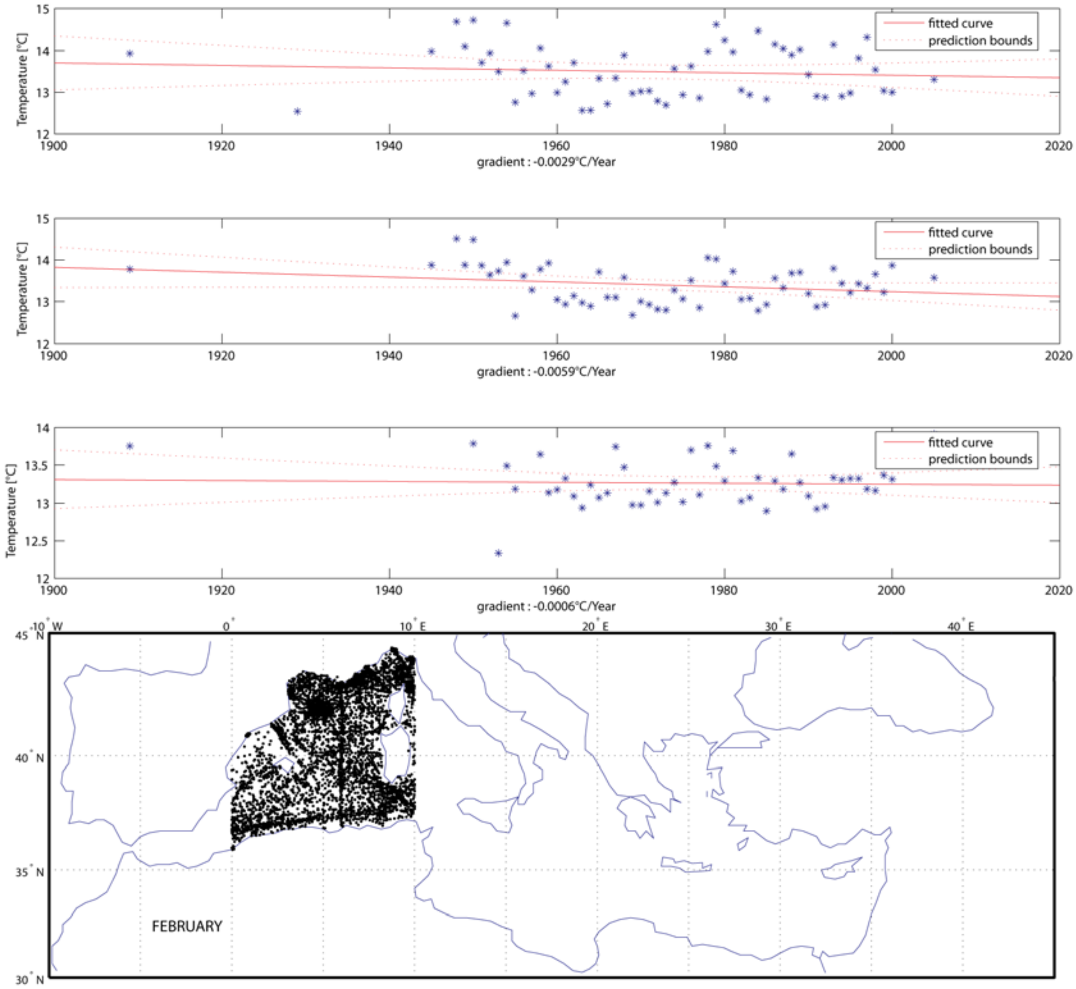


Figure 16a. The temporal variability of the temperature at surface, 300 m and 800 m during the month of February in the western Mediterranean.

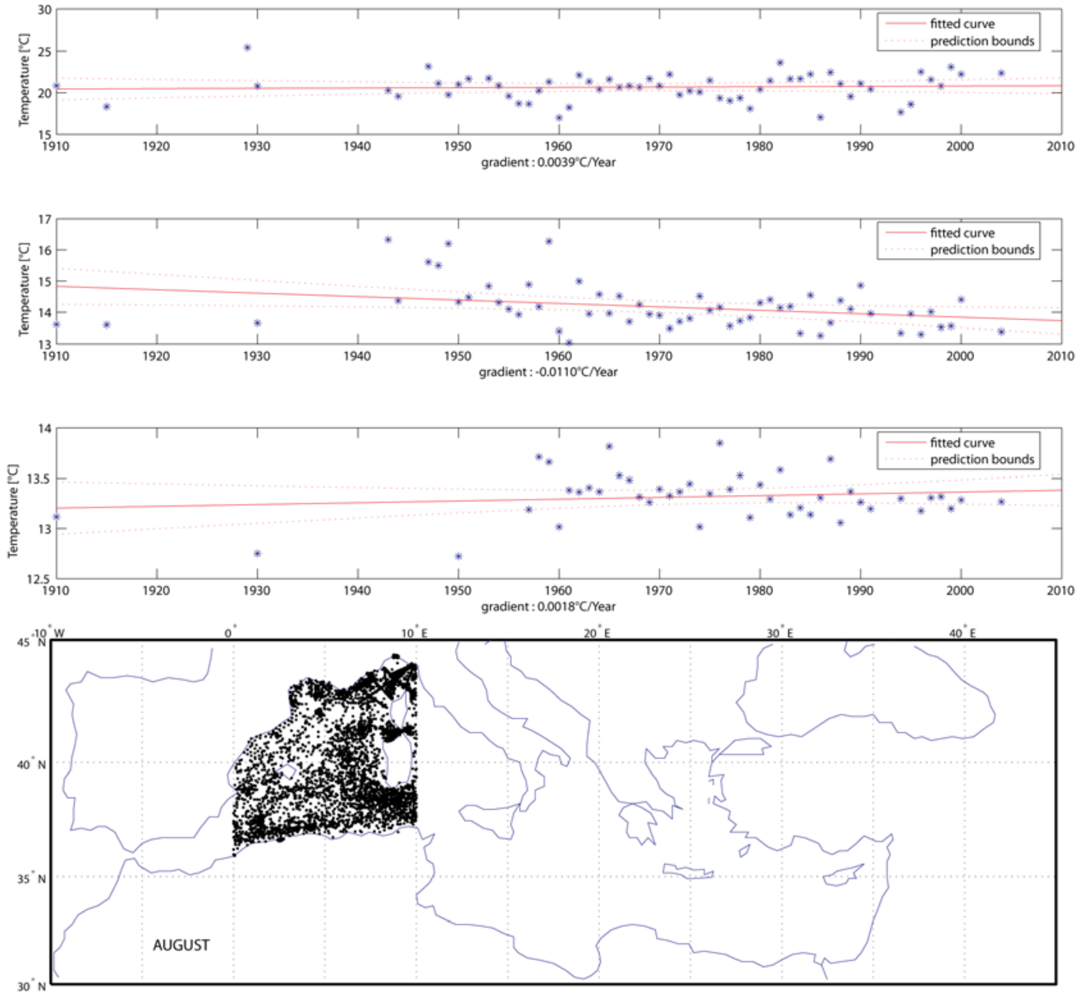


Figure 16b. The temporal variability of the temperature at surface, 300 m and 800 m during the month of August in the western Mediterranean.

6 New steps for quality control

A near real time quality control of temperature profiles collected with ships of opportunity was implemented in the Mediterranean during the Mediterranean Forecasting System – Pilot Project (Manzella et al., 2003). During the last year a significant amount of data were collected, and there is now the possibility to better specify the temporal and spatial variability of temperature. Furthermore, it is now possible to implement in a more sophisticated way the quality control procedures by using new reference profiles to be used for comparison with data to be controlled. The near real time control for XBT derived temperature profiles actually applied in MFS is consisting in 7 steps that in synthesis are:

1. position control,
2. elimination of spikes,
3. interpolation at 1m interval,
4. Gaussian smoothing,
5. general malfunction control,
6. gross range check
7. comparison with climatology,
8. visual check, confirming the validity of profiles and providing an overall consistency.

During the last years the original Ships Of Opportunity Programme in the Mediterranean has been complemented with a so called ‘Cruises Of Opportunity Programme’, i.e. research vessels were used to collect data to be used in near real time. This has required the implementation of the original quality control procedures, since also CTD data have been transmitted in near real time. As part of MyOcean project, we have substituted the temperature climatology with a best estimate of profiles, as derived from a statistical distribution of data in 1x1 degrees boxes. Furthermore, we have enlarged the near real time quality control also to other parameters (e.g. salinity, oxygen, nutrients). The gross range check is done in sub-regions shown in figure 17.

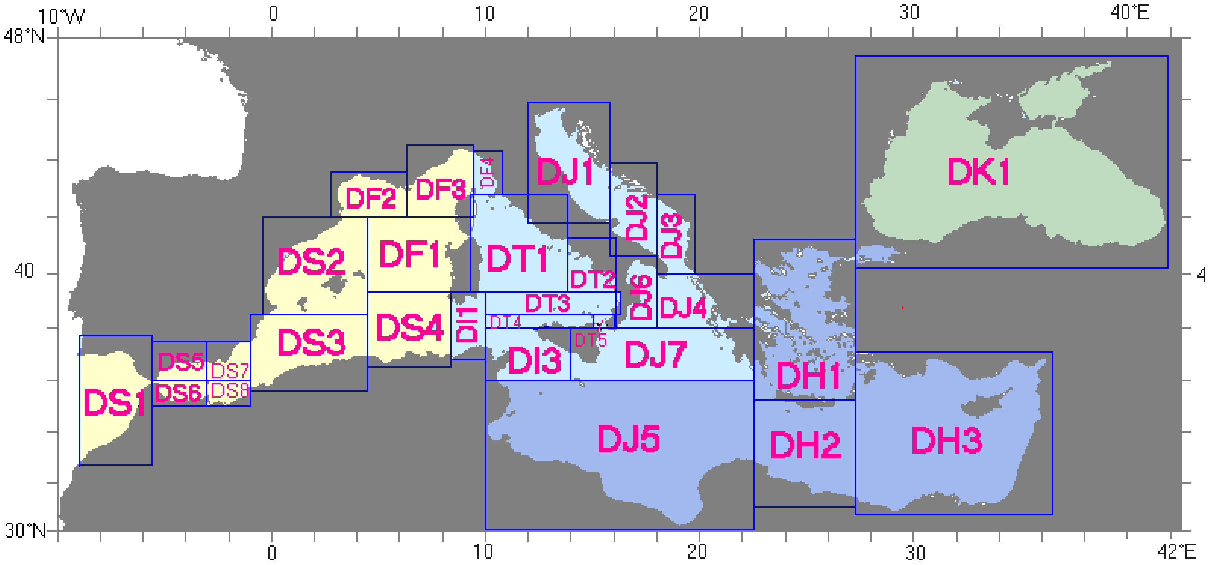


Figure 17. Areas for regional checks. In each subregion is defined the minimum and maximum value of different parameters.

For the temperature parameter we have defined best estimates of profiles in 1x1 degrees boxes and the temporal variability of the temperature in the different Mediterranean regions.

In the new procedure, the steps 5, 6 and 7 of the MFS quality control procedures have been modified and the overall consistency can take into consideration the existence of other dynamical states in some areas and the long term temporal variability (e.g. the temperature profile derived from second maximum in the statistical distribution). This kind of information can be added into meta-files and provide more useful information for reanalysis and studies on climate variability. The quality control has been modified as follows:

1. date and position control (using the ETOPO1 bathymetry, for both near real time and historical data), as well as check of the ship speed,
2. elimination of spikes (for XBT),
3. interpolation at 1m interval (eventually using software provided with instruments),
4. Gaussian smoothing (for XBT),
5. general malfunction control ,
6. regional range check,
7. overall profile quality control,
8. comparison with best estimate,
9. comparison with historical data to assess the temporal variability,
10. property/property scatter to assess the consistency of controlled data with other historical values,
11. visual check, confirming the final validity of profiles.

The step 7 provide the metadata with a flag indicating the quality of the entire profile. The general qualitative assessment of step 9 include also the evaluation of the technology used for data acquisition. A comment is added to the metafile accompanying the data. It must be underlined that this qualitative assessment is strongly based on experience of the data manager.

The procedures have been automated and are applied both to near real time and historical data. The software has been developed in fortran language and the visualization is done with Ocean Data View. An example is provided by using data collected by CNR/ISMAR/SP in the Strait of Sicily during November 2009 on board the RV Urania. These data include temperature, salinity, turbidity and fluorimetry and are treated as 'near real time'. As an example in table 3 is provided the output of the software checking the ship speed.

STATION NAMES AND DATES						
STATION	**DAY**	**MONTH**	**YEAR**	**HOUR**	**MIN**	**SEC**
d107.cnv	**31**	**10**	**2009**	**15**	**16**	**36**
d25.cnv	**17**	**11**	**2009**	**16**	**22**	**33**
d26.cnv	**18**	**11**	**2009**	**1**	**50**	**21**
d48.cnv	**17**	**11**	**2009**	**14**	**49**	**31**
d49.cnv	**17**	**11**	**2009**	**13**	**38**	**43**
d52.cnv	**2**	**11**	**2009**	**16**	**50**	**41**
MAXIMUM SPEED OF THE SHIP = 10.00						
stat.1 name	**stat.2 name**	**distance(nm)**	**Time interval (hours)**	**ship speed**		

d107.cnv	**d52.cnv**	**122.99**	**49.57**	**2.48**	WARNING! TOO SLOW
d52.cnv	**d49.cnv**	**418.66**	**356.80**	**1.17**	WARNING! TOO SLOW
d49.cnv	**d48.cnv**	**10.01**	**1.18**	**8.45**	EXPECTED SPEED
d48.cnv	**d25.cnv**	**4.34**	**1.55**	**2.80**	EXPECTED SPEED
d25.cnv	**d26.cnv**	**8.70**	**9.47**	**0.92**	WARNING! TOO SLOW

Table 3. Output of the QC software. Ship speed check. The interpretation of the speed requires a logbook.

The warning must be evaluated on the base of the cruise logbook, since a long time interval can be due to bad weather, or other causes. The software is checking also the position (land or sea) and the depth of the sea to be compared with the last data value or depth given in the metadata. This is done using the ETOPO1 data set.

CHECK LAND POSITION AND DEPTH WITH ETOPO1 DB				
Station	**Latitude**	**Longitude**	**Cast depth**	
d49.cnv	**36.3697**	**15.2797**	**143.**	
d52.cnv	**41.0002**	**10.0598**	**987.**	
d107.cnv	**43.0290**	**9.7677**	**86.**	CAST DEEPER THAN DEPTH d107.cnv 82.
d25.cnv	**36.5363**	**15.4498**	**1077.**	CAST DEEPER THAN DEPTH d25.cnv 1051.
d26.cnv	**36.4943**	**15.5887**	**1909.**	CAST DEEPER THAN DEPTH d26.cnv 1819.
d48.cnv	**36.4797**	**15.4050**	136.	CAST DEEPER THAN DEPTH d48.cnv 119

Table 4. Output of the QC software. Depth and land check. The software is warning the managed that two casts were too deep with respect to ETOPO1 data base. This requires an additional check with bathymetric charts to find that the casts were done on the Malta escarpment, where the 1 prime resolution is inadequate.

Than regional check and comparison with best estimate is done and results viewed with an ODV software. In this way it is possible to compare the profile to be controlled not only with the closest grid point of the best estimates, but also with the neighbours points (figure 18).

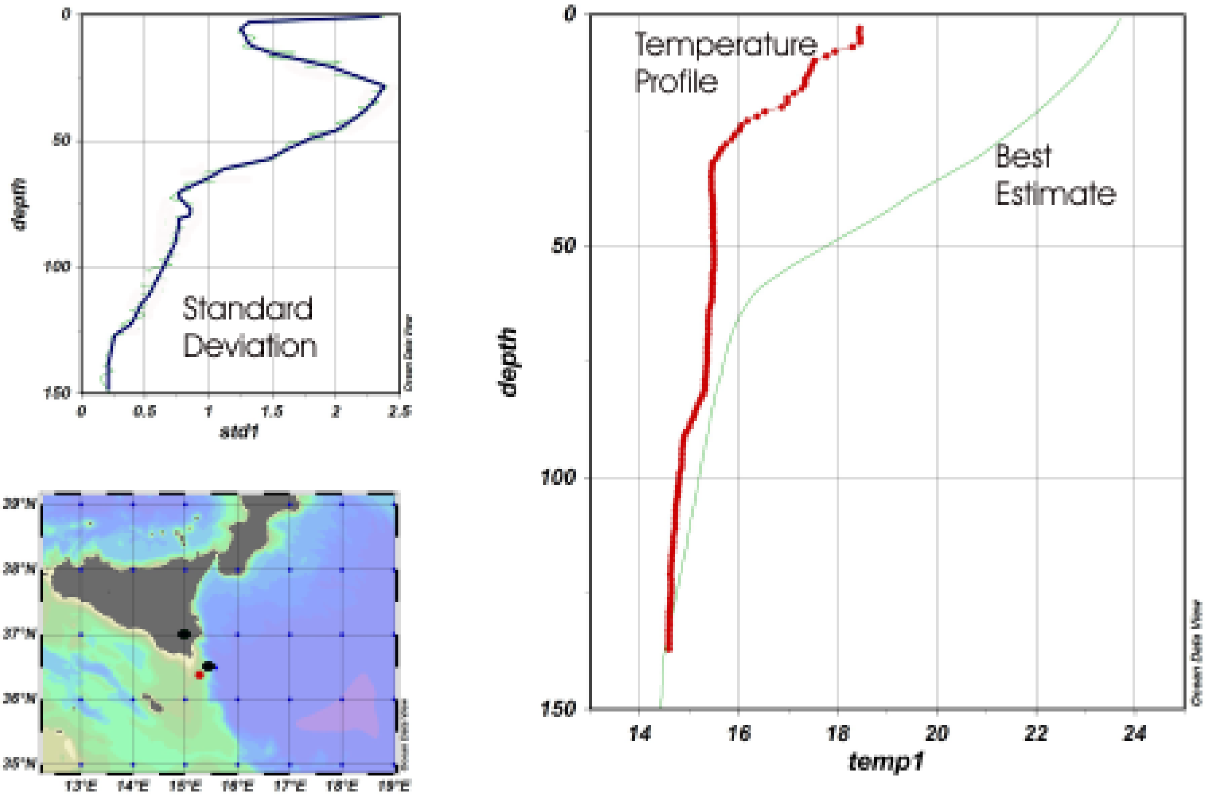


Figure 18. The temperature profile under the quality control process, the closest best estimate profile and the standard deviation profile. Temperature values are always within 3 standard deviations from best estimates. Clicking on different grid points it is possible to compare the profile with other best estimates.

Mean values in different regions have been calculated for each month in the different years in order to have an estimate of temporal variability. A minimum of 25 profiles was required to calculate these means, producing graphs (figures 13 –16) and tables. From these it is possible to assess how the data contribute to the temporal variability. This is, of course, a qualitative assessment, that can be included as a comment in metafiles.

November – Central Mediterranean			
Year	**Mean Temperature**	**Standard Deviat.**	**n.values**
1996	**19.259**	**0.862**	**1300**
1997	**16.944**	**1.657**	**750**
1998	**21.331**	**1.384**	**1000**
1999	**20.155**	**1.699**	**11350**
2000	**19.016**	**0.192**	**900**
2001	**20.013**	**2.507**	**1450**
2002	**18.689**	**0.781**	**1250**
2003	**No data**	**No data**	**No data**
2004	**21.587**	**2.146**	**5400**
2005	**20.291**	**1.946**	**5550**
2006	**19.147**	**1.915**	**600**
2009	18.100	-	Station D48

Table 5. Comparison of the temperature in the upper 50 m at station d48 with mean historical value in the central Mediterranean (Strait of Sicily).

Finally, it is possible to include also a property/property analysis, again using the ODV software (figure 19). Pre-requisite for this is the availability of historical data base in ODV to which the new data are added. Using the facilities provided by the software it is possible to complete the quality control, without significant differences to the delayed mode procedures.

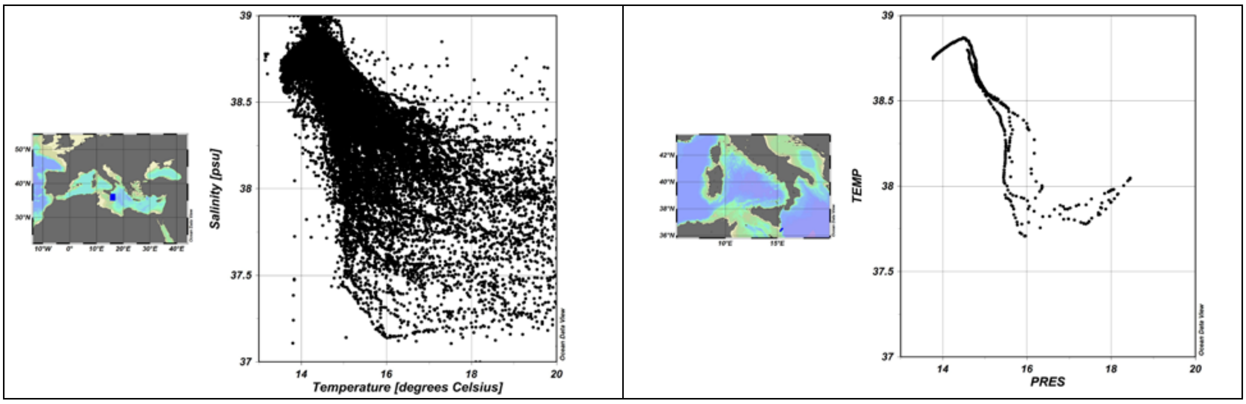


Figure 19. The T/S diagrams of historical data (left) and new data (right). The comparison allow the final assessment of the data.

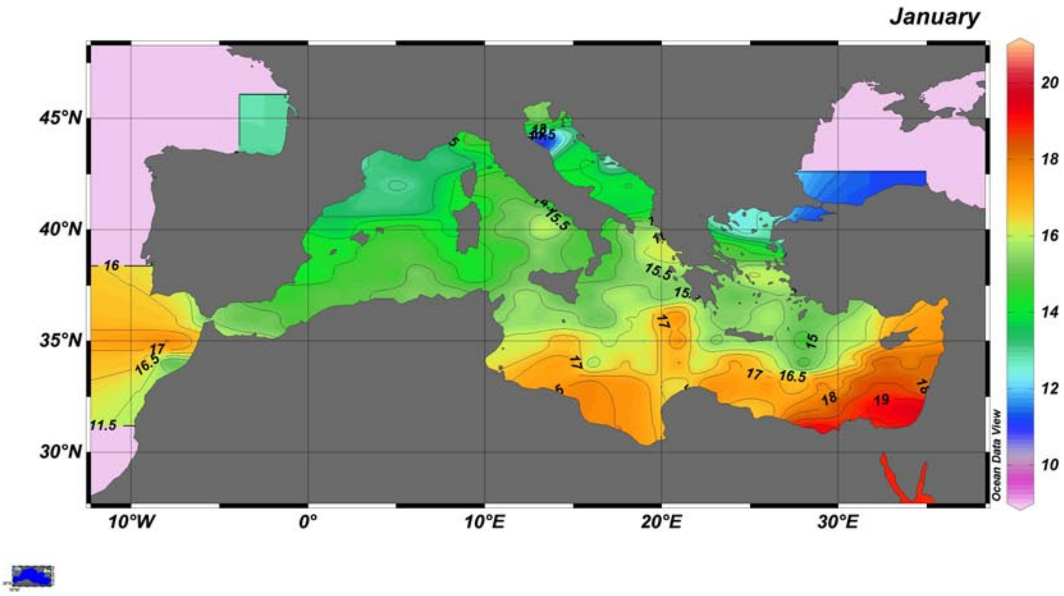
The instrument used for the data collection was a SeaBird900, i.e. the associated precision is quite high and no additional sources of errors can be considered.

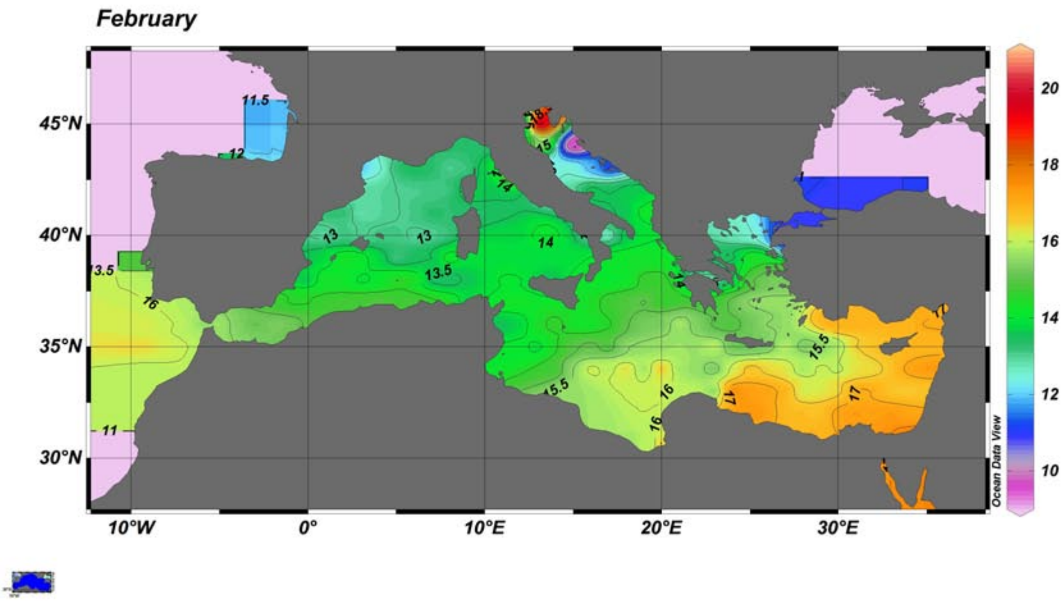
7 Conclusions

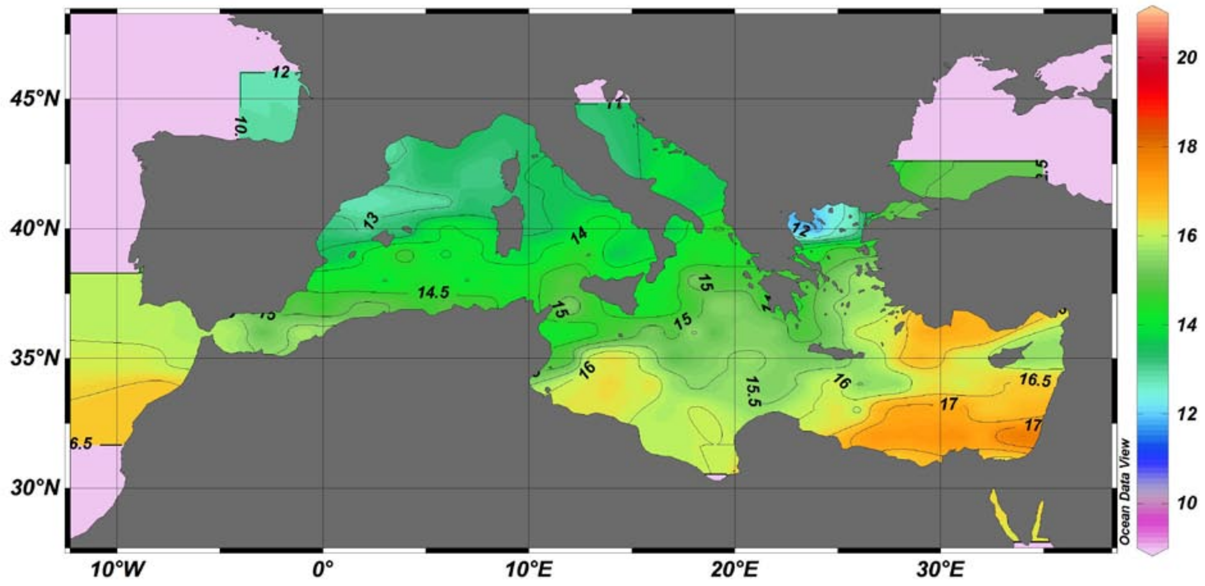
The exercise presented in this report demonstrate that quality control of near real time and delayed mode data can be done with identical procedures. This is an important improvement for an operational observing system based on multiplatforms, and then providing different type of parameters. The improved procedures takes few hours of an expert data manager for a significant amount of data, that is that the data can be released a short time after data collection, and in any case within the 12 hours defined as a target for the operational systems.

Some work is still needed to have a complete the new near real time procedures for the main parameters. At least best estimates for salinity must be obtained from historical data. However, all the tools are now ready.

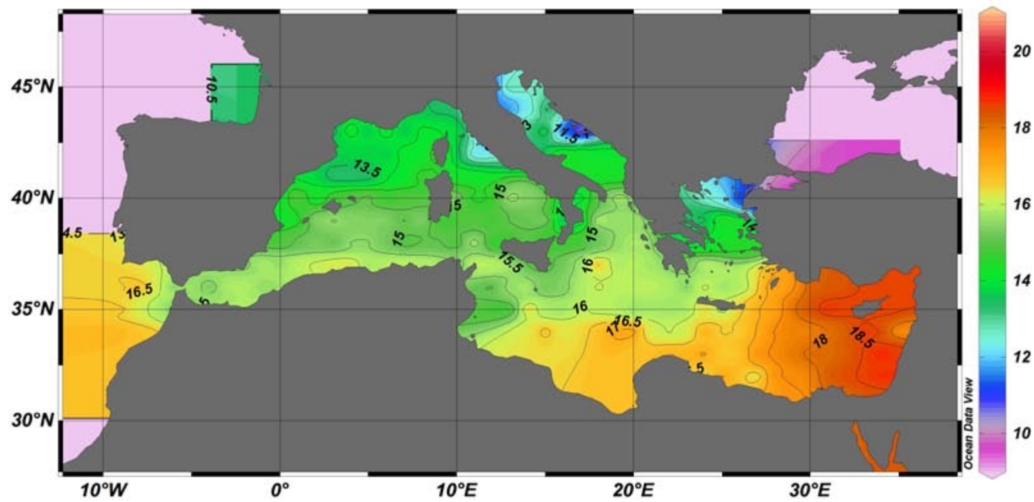
Appendix 1.
Surface maps of temperature from best estimates. The csale is maintained the same for all maps, the contour interval is 0.5 °C



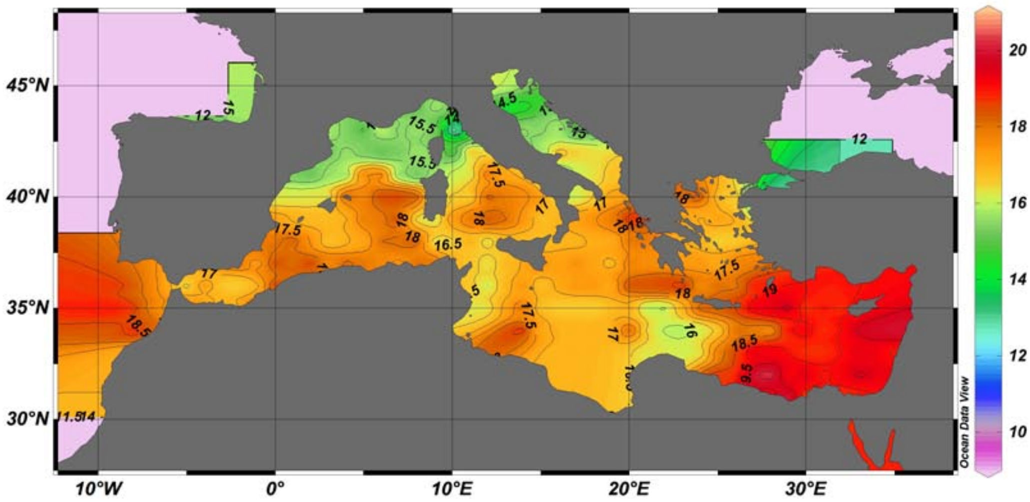



45°N
40°N
35°N
30°N
10°W
0°
10°E
20°E
30°E
20
18
16
14
12
10
Ocean Data View

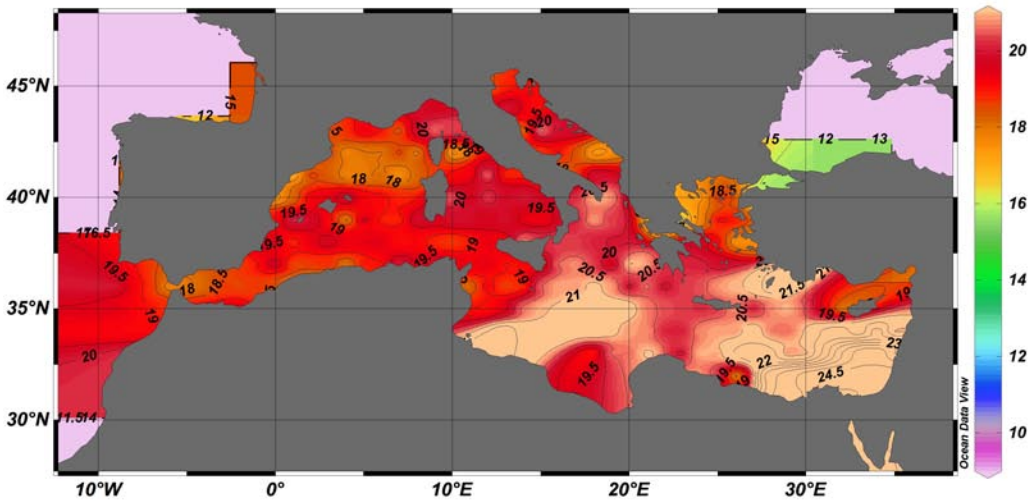
April


45°N
40°N
35°N
30°N
10°W
0°
10°E
20°E
30°E
20
18
16
14
12
10
Ocean Data View

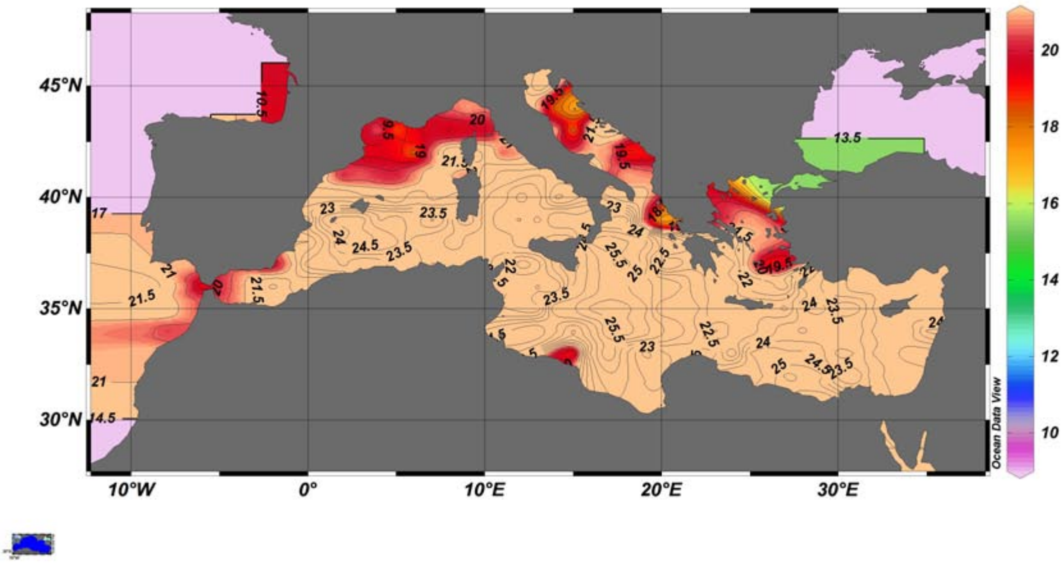
May


45°N
40°N
35°N
30°N
10°W
0°
10°E
20°E
30°E
20
18
16
14
12
10
Ocean Data View

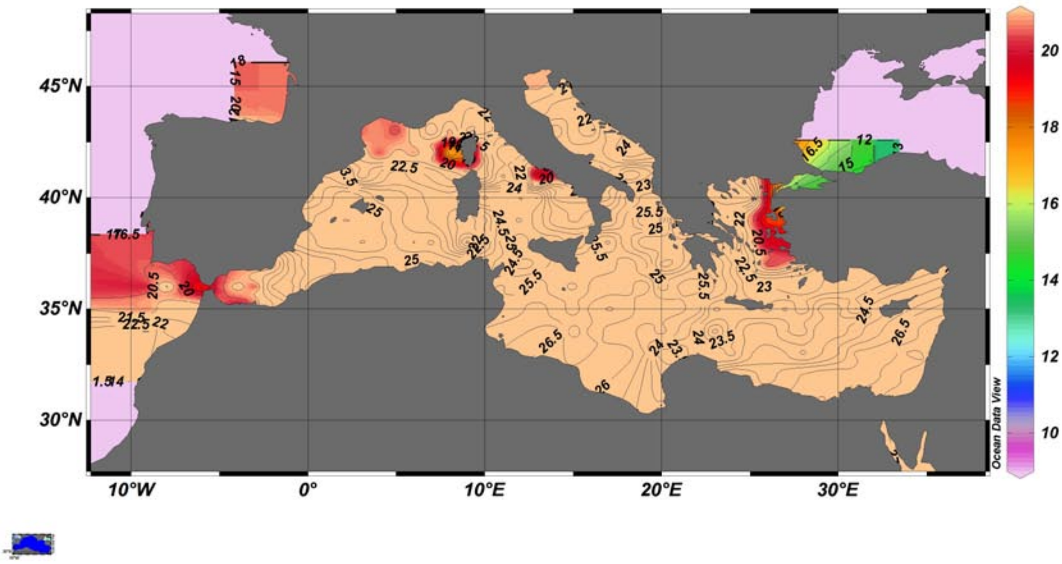
June


45°N
40°N
35°N
30°N
10°W
0°
10°E
20°E
30°E
20
18
16
14
12
10
Ocean Data View

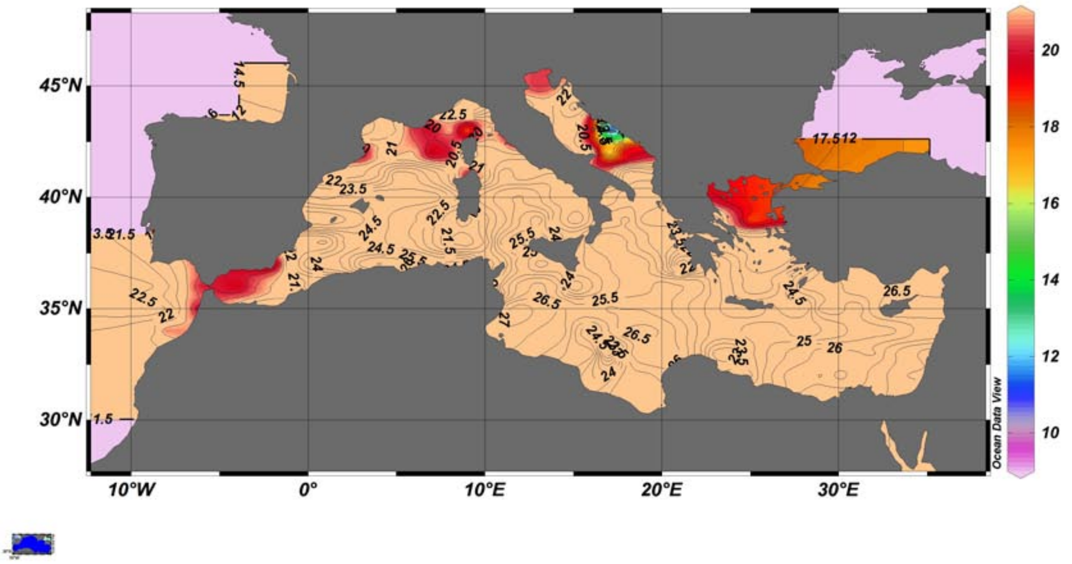
July


45°N
40°N
35°N
30°N
10°W
0°
10°E
20°E
30°E
20
18
16
14
12
10
Ocean Data View

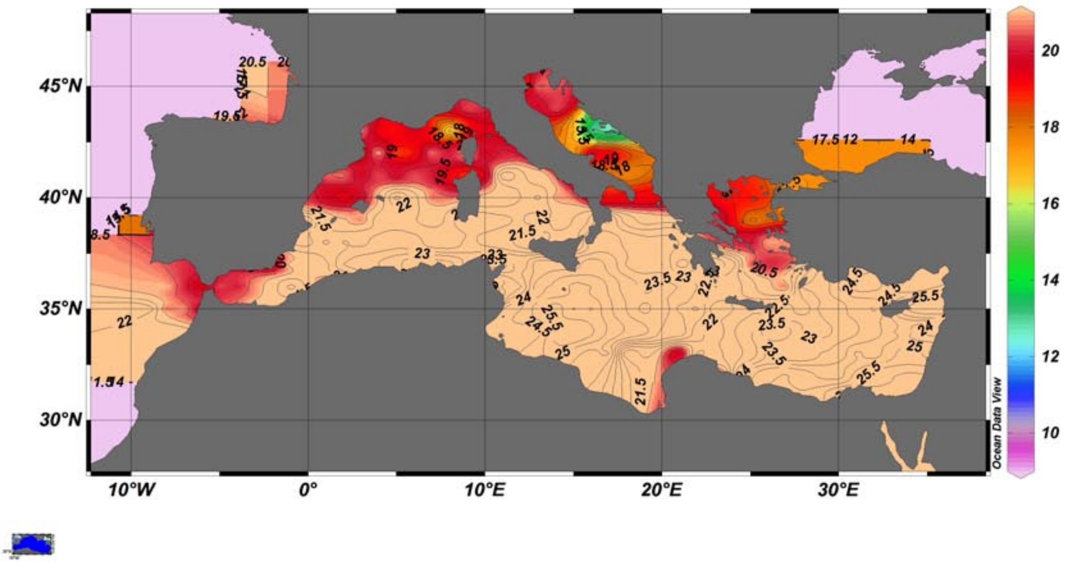
August

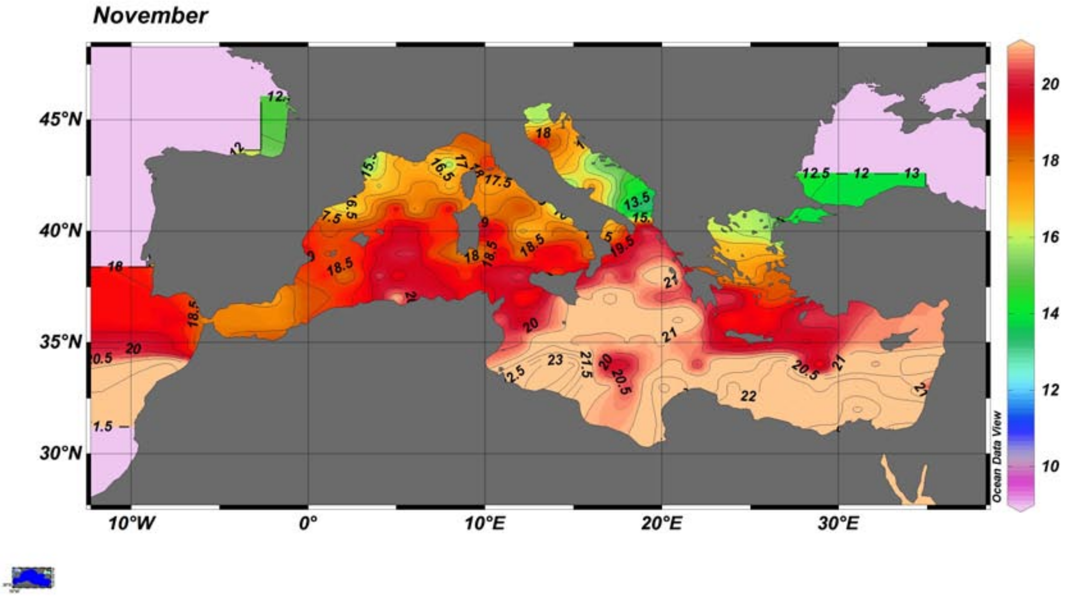

45°N
40°N
35°N
30°N
10°W
0°
10°E
20°E
30°E
20
18
16
14
12
10
Ocean Data View

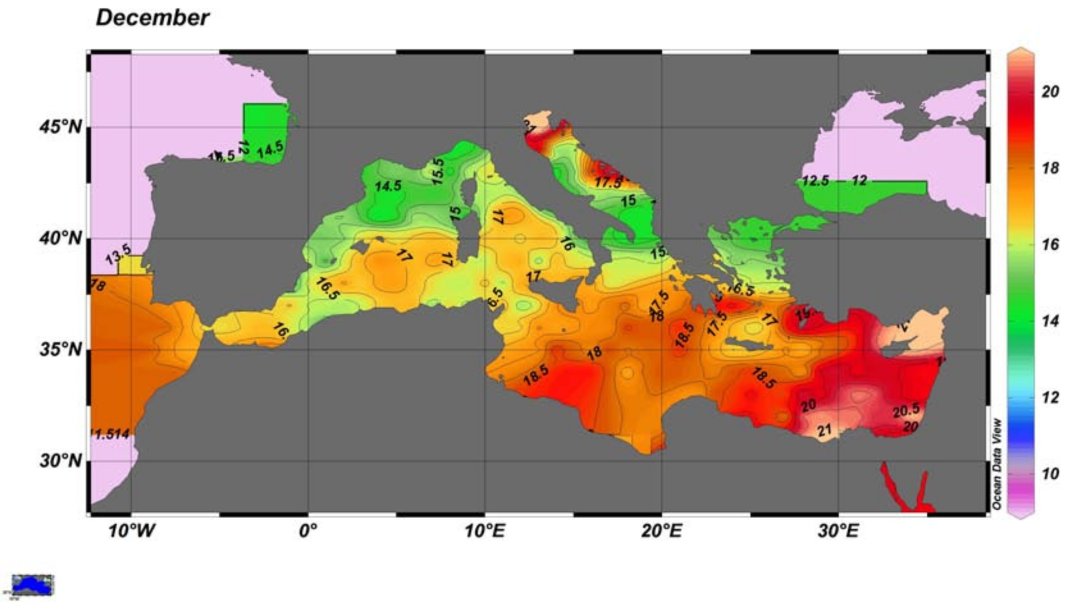
September



October




November
45°N
40°N
35°N
30°N
10°W
0°
10°E
20°E
30°E
20
18
16
14
12
10
Ocean Data View


December
45°N
40°N
35°N
30°N
10°W
0°
10°E
20°E
30°E
20
18
16
14
12
10
Ocean Data View

References

J.C. Davis (1973) Statistics and data analysis in geology. Wiley, XIII+550p.

V. Gouretski, K.P. Koltermann (2007) How much is the ocean really warming? Geophysical Research Letters, 34, L01610, 5 PP., 2007, doi:10.1029/2006GL027834

M.R. Howe, R.I. Tait (1965) An evaluation of an in-situ salinity-temperature-depth measuring system. Marine Geology 3, 6, 483-487, doi:10.1016/0025-3227(65)90048-4.

H. Lacombe, P. Tchernia, L. Gamberoni (1985). Variable bottom water in the Western Mediterranean Basin. Progress in Oceanography, 14, 319-338.

S. Levitus, J. I. Antonov, T. P. Boyer, C. Stephens (2000). Warming of the World Ocean. Science 24 March 2000: Vol. 287. no. 5461, pp. 2225 – 2229 DOI : 10.1126 /science.287.5461.2225

C. Maillard, M. Fichaut and MEDAR/MEDATLAS GROUP (2001) Medar-Medatlas protocol. R.INT.TMSI/IDM/SISMER/SIS00-084

G.M.R. Manzella, E. Scoccimarro, N. Pinardi, M. Tonani. Improved near real-time data management procedures for the Mediterranean ocean Forecasting System – Volunteer Observing Ship program. Annales Geophysicae, 21, 49-62, 2003

MEDAR Group (2002), Mediterranean and Black Sea Database of Temperature, Salinity and Biochemical Parameters and Climatological Atlas [4 CD-ROMs], Ifremer Ed., Plouzane, France. (Available at http://www.ifremer.fr/sismer/program/medar/)

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

N. Pinardi, G. Coppini (2007). Operational oceanography in the Mediterranean Sea: the second stage of development. Ocean Sciences, Ocean Sci., 6, 263–267, 2010 www.ocean-sci.net/6/263/2010/

M. Rixen, J.-M. Beckers, S. Levitus, J. Antonov, T. Boyer, C. Maillard,M. Fichaut, E. Balopoulos, S. Iona, H. Dooley, M.-J. Garcia, B. Manca, A. Giorgetti, G. Manzella, N. Mikhailov, N. Pinardi, and M. Zavatarelli (2005), The Western Mediterranean Deep Water: A proxy for climate change, Geophysical Research Letters, Vol. 32, L12608, doi:10.1029/2005GL022702

A.R. Robinson, M. Golnaraghi, W.G. Leslie, A. Artegiani, A. Hecht, E. Lazzoni, A. Michelato, E. Sansone, A. Theocharis and Ü. Ünlüata (1991). The eastern Mediterranean general circulation: features, structure and variability. Dynamics of Atmospheres and Oceans, 15, 215-240.

WMCE Consortium of Investigators (1989). Western Mediterranean Circulation Experiment: a preliminary review of results. EOS, 70, 31.