SDC_V1 Release of Climatologies and Documentation for regional seas and the global ocean

WP11 - Deliverable D11.6

HORIZON **2020**

sdn-userdesk@seadatanet.org – www.seadatanet.org

SeaDataCloud - Further developing the pan-European infrastructure for marine and ocean data management

Grant Agreement Number: 730960

Deliverable number	Short title
D11.6	First release of SDC Climatologies

Long title

SDC_ V1 Release of Climatologies and Documentation for regional seas and the global ocean

Short description

The report summarizes the methodology applied to compute the regional and global climatologies; it describes main products' characteristics, their consistency analysis results.

Author	Working group
S. Simoncelli, A. Barth, C. Troupin , V. Myroshnychenko and P. Oliveri with WP11 team	WP11

Dissemination	Copyright terms
Public	

History

Version	Authors	Date	Comments
V0	S.Simoncelli ,A. Barth, C. Troupin	May 2019	outline definition and first draft
V1	S.Simoncelli	24 July 2019	first version
V2	V. Myroshnychenko and P. Oliveri	24 July 2019	input in some sections
V3	M. Fichaut	25 July 2019	Final review

Disclaimer

Acknowledgment

This document is a deliverable of the SeaDataCloud project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 730960.

Table of contents

1. Introduction

The objectives of Task 2 (Produce Standard Climatological Data Products) were defined in the deliverable D11.5 and hereafter reported:

1. To produce Temperature and Salinity climatologies with increased horizontal and vertical resolution with respect to SDN2 ones;
2. To compute decadal or pentadal climatologies depending on the data distribution in time and space together with their long-term variability;
3. To perform consistency analysis of SDC climatologies with available World Ocean Atlas (WOA) climatologies and/or other products like CMEMS climatologies computed by satellite reprocessed data sets and reanalysis products; EN4 gridded products; MEDAR-MEDATLAS climatology
4. To include all available SDC restricted observations and integrate external data sets in order to provide decadal climatologies.

During the feasibility analysis of new climatological products, a review of World Ocean Atlas V2 2013 (Locarnini et al., 2013 and Zweng et al., 2013) has been crucial to understand what could have been done at regional and global scales to progress with respect to the products already available. In the meantime a new WOA 2018 (Locarnini et al., 2018 and Zweng et al., 2018)has been released and this allowed the regional leaders to consider the 2013 V2 or the latest version available.

The decision to conform to WOA schema to produce SDC climatologies and adopt their vertical standard levels (see Figure 1) was crucial since a new DIVAnd software has been used and a common consistency analysis with WOA fields guarantees its good performance.

The Table in Appendix 1 summarizes the SeaDataCloud climatologies and their characteristics, that have been produced after a detailed analysis of data spatial and temporal distribution in each domain.

A global SDC product (SDC_GLO_CLIM_TS_V1) has been created for the first time which contains two different monthly climatology for temperature and salinity, SDC_GLO_CLIM_TS_V1_1 and SDC_GLO_CLIM_TS_V1_2 with a different time coverage, computed from data from the World Ocean Database (WOD). This choice has been taken because spatial coverage of SeaDataNet data at global scale is still too sparse, but in the next releases the idea is to integrate both data sources and improve at best the product's quality.

Regional climatologies were designed with a harmonized initial approach and all cover the time period after 1955, when marine data start to be sufficient for mapping at regional scale. All regional products are characterized by a monthly climatological field covering the entire time span 1955-2014 at least (some of them reaches 2017) and decadal climatologies at seasonal temporal resolution (monthly for the BAL region). Moreover all of them have been created integrating, for the first time, SDC aggregated datasets with external sources (World Ocean Database and COriolis Ocean Dataset for Reanalysis), which highly increased the spatial data coverage.

All climatological products and their generation methodology and validation have been described in their relative Product Information Documents (PIDoc), which is available together with the NetCDF files at the Sextant Catalogue. Their publication is still on going but it will be soon finalized as detailed in deliverable D11.12. Each PIDoc passed through a double stage revision by the WP leader and ULiege team that developed DIVAnd software, in order to assure a high quality documentation and increase user confidence and uptake.

The present report aims at summarizing the main phases of the climatologies' generation and the differences/commonalities, strength/weakness adopted at regional level. The report outline follows the PIDoc one: the description of the input data sets; the external data integration; the methodology and DIVA implementation; the results; the consistency analysis and the concluding remarks.

Depth (m)	Level	Depth (m)	Level	Depth (m)	Level	Depth (m)	Level
0	1	475	36	2300	70	5700	104
5	2	500	37	2400	71	5800	105
10	3	550	38	2500	72	5900	106
15	4	600	39	2600	73	6000	107
20	5	650	40	2700	74	6100	108
25	6	700	41	2800	75	6200	109
30	7	750	42	2900	76	6300	110
35	8	800	43	3000	77	6400	111
40	9	850	44	3100	78	6500	112
45	10	900	45	3200	79	6600	113
50	11	950	46	3300	80	6700	114
55	12	1000	47	3400	81	6800	115
60	13	1050	48	3500	82	6900	116
65	14	1100	49	3600	83	7000	117
70	15	1150	50	3700	84	7100	118
75	16	1200	51	3800	85	7200	119
80	17	1250	52	3900	86	7300	120
85	18	1300	53	4000	87	7400	121
90	19	1350	54	4100	88	7500	122
95	20	1400	55	4200	89	7600	123
100	21	1450	56	4300	90	7700	124
125	22	1500	57	4400	91	7800	125
150	23	1550	58	4500	92	7900	126
175	24	1600	59	4600	93	8000	127
200	25	1650	60	4700	94	8100	128
225	26	1700	61	4800	95	8200	129
250	27	1750	62	4900	96	8300	130
275	28	1800	63	5000	97	8400	131
300	29	1850	64	5100	98	8500	132
325	30	1900	65	5200	99	8600	133
350	31	1950	66	5300	100	8700	134
375	32	2000	67	5400	101	8800	135
400	33	2100	68	5500	102	8900	136
425	34	2200	69	5600	103	9000	137
450	35						

Figure 1 - Table summarizing the depth associated with each standard level number (Table 3 from WORLD OCEAN ATLAS 2018 Product Documentation).

2. Description of input data set

2.1. Description of input data sets

Global climatology is the only one produced using the single data source WOD 2013 V2 (Boyer et al., 2013), while all regional climatologies have been created integrating, for the first time, the SDC aggregated datasets with external sources (World Ocean Database and COriolis Ocean Dataset for Reanalysis). This highly increased the spatial data coverage thus improving the products' quality. SDC data both restricted and unrestricted, passed through a detailed quality control procedure. Only SDC data flagged as good (QF=1) general considerationsand probably good (QF=2) were selected for the analysis.

The COriolis Ocean Dataset for Reanalysis (CORA) dataset is an extraction of the Coriolis database at a given date each year and it is available from the COPERNICUS Marine and Environment Monitoring Service (CMEMS) (http://marine.copernicus.eu/services-portfolio/ product's code INSITU_GLO_TS_REP_OBSERVATIONS_013_001_b). The CORA dataset version 5.1 or version 5.2 (Szekely et al., 2016 and 2019) has been integrated into the SeaDataCloud data sets to produce climatologies (see Table 1). The extraction of the NetCDF files has been done through a common procedure and it took into account the QC on the date and the position (Quality Flag =1 good) and exported adjusted parameters when they exist.

World Ocean Database data have been downloaded (https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html) and integrated as well in its 2013 (Boyer et al., 2013) or 2018 version (Boyer et al., 2018) release (see Table 1).

Table 1 reports a synthesis of the amount of data used per each data source to construct the climatology data sets per each sea region.

The first global climatology (SDC_GLO_CLIM_V1_1) is estimated using profiles from Conductivity Depth Temperatures CTD, Profiling Floats (PFL), Ocean Station Data (OSD) and Moored buoy data (MRB) from 1900 to 2017. The second global climatology (SDC_GLO_CLIM_V1_2) is estimated only using the (PFL) data set from 2003 to 2017.

The Arctic Ocean data set presents a low SDN data availability in the first 4 decades, while for the last two decades the data availability is comparable to WOD18 one. This explains the huge difference in the number of data of the two sources.

The North Sea data set present a good balance between SDC and WOA18 amount of stations considered.

The Baltic Sea SDC data set has been merged with CORA5.2, which constitutes 35% of the whole climatology input data set.

The North Atlantic data set has been integrated with CORA5.1 to address the climatology, only some data types were selected:

- CT: CTD data from research vessels but also data from sea mammals equipped with CTD and some Sea Gliders.
- OC: CTD and XCTD coming from the high resolution CTD dataset of the World Ocean Database
- ME: CTD from Sismer database, coming from the French oceanographic campaigns.
- IC: CTD from ICES dataset, those profiles complete the CTDs coverage in CORA on the period 1900-2011.
- OS: OceanSites data that are mostly CTD (Oceansites moorings are in TS_MO)
- PF: data from Argo floats directly received from the Argo DACs.
- SH: data from the SHOM database (most of them cover the period 1950-1990 period).

The number of stations have not been reported in PIDoc but provided as graphs but approximate statistics on the percentage of stations are estimated and it results that SDC data cover about the 67% of the climatology data set, while CORA5.1 the 33%.

The Mediterranean data set integrates SDC unrestricted and restricted data with CORA5.2. Only Temperature and Salinity couples and none ferrybox data have been considered to produce the climatological fields to facilitate consistency among Temperature and Salinity. The same CORA data types of the North Atlantic product have been selected except, the OC one which contained a lot of stations on land but flagged as good.

The Black Sea data set is the only one that blends SDN, WOD18 and CORA5.1, all the main marine data sources, and it consists of 57% of SDN data (both restricted and unrestricted), 20% of data from WOD18 and 23% from CORA5.1.

Table 1 - Input data sets used to produce SDC_CLIM_TS_V1 and the relative number of stations. * Baltic statistics are provided as number of measurements and not stations.

Name	N. of stations SDC unrestricted data	N. of stations SDC restricted data	External data set	N. of stations External data set	
SDC_GLO_CLIM_TS_V1_1			WOD13 (PFL, OSD, MRB)	7039304	
SDC_GLO_CLIM_TS_V1_2			WOD13 (PFL)	1862686	
SDC_ARC_CLIM_TS_V1	148044 (24%) without ferrybox	373	WOD18 (PFL, OSD)	479118 (76%)	627535
SDC_NS_CLIM_TS_V1	162452 (54%)		WOD18 (PFL, OSD)	168197 56%	301238
SDC_BAL_CLIM_TS_V1	13727702* (65%)	121526* (<1%)	CORA5.2	7331275* (35%)	21180503*
SDC_NAT_CLIM_TS_V1	9421968 (~67%)		CORA5.1	(N not available) (~33%)	(N not available)
SDC_MED_CLIM_TS_V1	143871 (61%) without ferrybox		CORA5.2	90883 (39%)	234754
SDC_BLS_CLIM_TS_V1	130466 (53%)	10285 (4%)	WOD18 CORA5.1	48227 (20%) 57847(23%)	246825

2.2. Additional Quality on external data sources

External sources' data have been imported to ODV (https://odv.awi.de/) as collection and additional quality check that has been performed through visual inspection in ODV, following the SDC QC guidelines detailed in D11.3. This step was necessary since all external data sources in Table 1 presented data anomalies, meaning data flagged as good but visibly wrong. This fundamental step increases data consistency, since each initiative apply different quality check strategies and procedures. It also avoids to deteriorate SDC product quality too.

Datasets integration was performed in general through the following steps:

- Excluding internal duplicates

- Identifying and excluding overlapping data
- Merging non-overlapping data

The steps of the merging procedure:

- The SDC data sets were taken as a primary;
- The SDC restricted data sets were added;
- Non-overlapping part of the external data sets were added.

The first step of QC process has been the duplicate removal, which has been performed in different ways:

- Through ODV (NAT and BLS regions) datasets have been checked with the algorithm provided by ODV and its standard parameters (< 0.001° in latitude and longitude, < 1h between the recorded times of measurement). In BLS dataset the found duplicate candidates then were tested for data matching and only confirmed matches were accepted as duplicates.;
- Through DIVAnd (NS and MED regions): duplicate detection and removal was performed using the tool DIVAnd.Quadtrees.checkduplicates, with the following values for the parameters (< 0.001° in latitude and longitude, < 2h between the recorded times of measurement, < 1m for depth, < 0.1°C for Temperature, <0.1 PSU for Salinity) [https://github.com/gher-ulg/SeaDataCloud/blob/master/Julia/Climatologies/duplicate_detection.jl]. The check for duplicates is thus done at the data level, not at the station level as done with ODV in the inspection of the data sources described above. The criteria on location and time difference are also wider than the standard values used by ODV.
- IMR duplicated removal tool for the ARC region;

In BAL region duplicates were not removed but the weighting option in DIVAnd has been applied, which assigns a lower weight to observations that are close in space and time. Doing so, the influence of data duplicates in the resulting climatologies should be insignificant.

Table 2 - Number of duplicates with respect to SDC data collections in each sea region.

Sea Region	WOD	CORA
ARC		
NS	2.4% measurements (DIVAnd) 2.6% stations (ODV)	
BAL		
NAT		
MED		14% measurements (DIVAnd) 16.8% stations (ODV)
BLS	46515 (16.7%)	8698 (3,3%) + 25136 with WOD (9.2%)

A successive QC analysis on the merged dataset has been applied, first through visual inspection in ODV, then removing data that generated artificial bull-eyes features in the gridded fields. These artificial features could be due to non-detected outliers in previous QC phases or to representativeness errors, meaning observations that do not represent the climatological status of the region under analysis, but rather smaller spatial and temporal scales' events.

This latest procedure can be carried out manually or using the DIVAnd tool which computes the observation residuals after the climatological fields generation. The residuals have been analysed and observations with residuals larger than pre-defined thresholds (number of standard deviations from the mean have been removed. This latest approach has been applied in the Black Sea and Mediterranean Sea analyses. This second order QC procedure has been implemented in the global domain as well but not applied to this version of gridded climatologies yet.

3. DIVA implementation and settings

Computation of SDC Temperature and Salinity climatic fields was done with DIVAnd (*Barth et al., 2014*). DIVAnd has been implemented in the programming language Julia (https://github.com/gher-ulg/DIVAnd.jl) and is used in conjunction with the Jupyter notebooks (https://jupyter.org/) – the web-based interactive computational environment for creating and sharing documents that contain live code, equations, visualizations and narrative text. This is particularly convenient for climatology generation, because the input files, analysis parameters, visualisations and outputs can be defined directly in a notebook, and also the parameter tuning task is much easier.

Data Interpolating Variational analysis (DIVA) is a tool to estimate continuous field from in-situ (observational) data. It is based on Variational inverse method that works with minimization of the cost function allowing the choice of analysed field fitting at best the data set.

Table 3 reports the principal settings of DIVA implementation per each sea region, in particular the correlation length, the error variance (epsilon2), the background field, the *surfextend* option. The **correlation length** (Lc) gives an indication of the distance over which a given data point influences its neighbourhood. The **error variance** (epsilon2) is the error variance of the observations (normalized by the error variance of the background field). It can be a scalar, vector or a matrix depending upon the error covariance of the observation and correlation of the error and in case of a scalar it is the *inverse of the signal-to-noise ratio* [4]. The **background field** is the first guess to the analysis. In case of default background field, diva3D is using spatial mean of the observations. The **surfextend option** when set as 'true' activates the use additional layer above the surface in order to improve first layer solution. All climatologies except the MED have been generated using the GEBCO (30 sec) bathymetry to construct the analysis topography. The MED climatologies has been generated using the new GEBCO 2019 (15 sec) bathymetry.

Table 3 - Summary of DIVA implementation per each domain.

	DIVA version	Lc	Epsilon2	surfextend	Background
SDC_GLO_CLIM_TS_V1	DIVAnd 2.3.1	Lc(x,y)=200km	0.9	false	spatial mean
SDC_ARC_CLIM_TS_V1	*DIVA master 4.7.2	Lc(x,y)=1 deg	2.0 (0.5 signal to noise ratio)	-	seasonal 1955-2014
SDC_NS_CLIM_TS_V1	DIVAnd 2.3.1	100km	0.1	true	spatial mean
SDC_BAL_CLIM_TS_V1	DIVAnd 2.3.1	Lc(x,y)=120km Lc(z)=20m	1.0		seasonal 1955-2014 Lc(x,y)=240km Lc(z)=50m
SDC_NAT_CLIM_TS_V1	*DIVA master 4.7.2	Lc(x,y)=2 deg	1.0	-	spatial mean
SDC_MED_CLIM_TS_V1	DIVAnd 2.4.0 (branch Alex)	Lc(x,y)=2 deg Lc(z)=0	0.6	false	T monthly S annual 1955-2017 1955-1984 1985-2017 Lc(x,y)=10 deg

SDC_BLS_CLIM_TS_V1	DIVAnd 2.3.1	Lc(x,y)=150km Lc(z) varying with depth	T= 0.3an, 0.5bkg S=0.1	true	Lc(x,y)=200km Lc(z)= varying with depth

During the creation of SDC_CLIM_TS_V1 products the WP11 team highly contributed to the testing and consolidation of DIVAnd software, even if two over seven products have been generated still with the previous DIVA master version.

NAT regional leader (C. Coatanoan) tested DIVAnd but some problems occurred. The main blocking points encountered were:

- to make DIVAnd compute horizontal and vertical correlation lengths for each timestep (i.e. len=() when calling function diva3D). This point has been solved.
- to give some weights to observations (i.e. epsilon2 values) when computing horizontal and vertical correlation for each time step (i.e. through diva3D function). This point could be bypassed, Lc could be also assigned.

Test result with DIVAnd shows better results where the data coverage is not uniform (Figure 2).

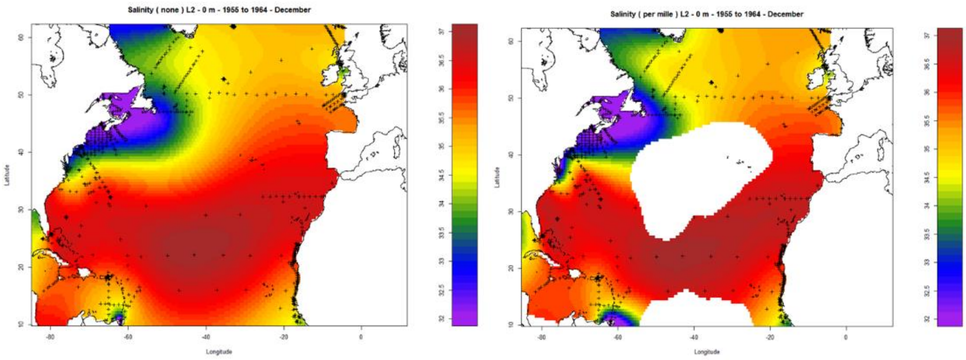


Figure 2 - Results for salinity in surface with (a) DIVAnd and (b) DIVA 4.7.2, for decade 1955-1964 in December.

GLO and MED climatologies used instead the DIVAnd in 2D mode, without exploiting the full potential of 3D analysis which is the novelty of DIVAnd tool. Main reason is the difficulty in tuning the vertical correlation length in domains where the variability is very high and the different water masses need to be well resolved along the water column.

Main topic of discussion among the WP11 team during the production period were:

- Background definition (see section 3.1)
- Correlation length tuning
- Vertical correlation length definition
- Vertical stability (see section 3.2)

It follows a summary of the main issues encountered and solved thanks to an intense team working.

Other issues:

- **Memory problems to handle the entire dataset**🖳 tuning of *memtofit* parameter: increase it memtofit to resolve runtime problems.
- **Matlab code for reading CORA files** (provided by BLS regional leader V. Myroshnychenko) call to *sw_dpth* function to convert pressure to depth. *sw_dpth* is from EOS-80 package. It takes into account the QF on position (*position_QC*) and date (*JULD_QC*). It considers the field *<parameter_name>_ADJUSTED*, when available, instead of the raw data.

- **Plotting the results**💻 *plotres* function has been improved
- **First surface layer:** A. Barth (ULiege) added the option *surfextend* to the function *diva3d*. When it is set to true, then an additional layer is placed on top such that the vertical resolution is the same between the first two layers. If you have vertical layers at a depth e.g. 0 m, 5 m, 10 m, then the additional layer is at -5 m. The additional layer is not saved in the output file. The additional layer should have the same land-sea mask as the very first layer in the water. Explicitly using the depth level at a negative value (outside of diva3d) does effectively use a mask who does not correspond to the land-sea mask because a negative depth value would represent a section in the air. A grid cell is considered land based on the average depth of a grid cell. So if you compare the data distribution with and without negative value of the depth vector, a difference is expected.

  The mask for the additional layer (z = 5) would be the same as for the actual layer (z = 10). The mask is quite important to define how the information is propagated spatially. If the make of the additional layer would be different than the first actual layer, then there would be some probably unexpected spatial connection.

  DIVAnd has been modified such that the additional layer is not passed to the plotting routine *plotres*
- **Changes in the way points near the coast line are handled:** all observations which are not surrounded by 4 sea grid points (or 8 in 3D dimensions) were no longer used. Now where a bi-linear interpolation is not possible, an interpolation to the nearest grid point is done: so among the 4 surrounding grid points (or 8 in 3D) only one grid point has to be a sea grid point to be retained. Previously 22% of data points where excluded for winter 1955-1964 in the spatial interpolation routine (after the time selection has been made). Now the discarded points reduced to 1.3%. To activate this statistics set ENV["JULIA_DEBUG"] = "DIVAnd" that gives as output the number of points outside 7374 or about 1.3%

  The parameter *sel* passed to *plotres* function reflects the actual data used to make the analysis at the particular time instance (after excluding points on "land" and negative values if an anamorphosis transform is used).
- **Isolated points along the coast**: using *DIVAnd.weight_RtimesOne* to reduce the weight also help to address this issue. However, this solution does not resolve the problem definitively since the "isolated points" are still present and still do not agree with the main field. in Black Sea such points were masked manually.
- **Land-sea mask creation:** since both the MED sea T and S fields have a great variability and the MED bathymetry is very complicated, the presence of wrongly water unconnected sub basins and of small water unconnected domains at the seafloor results in high gradients in analysed fields in these areas and in the immediate neighbourhood. In order to avoid this problem, a lot of manual edits in the land-sea mask have been implemented and the unconnected water areas at the seafloor with 3 or less points have been removed.
- **DIVAnd2ODV Matlab code for converting resulted .nc file** (provided by BLS regional leader V. Myroshnychenko) to ODV spreadsheet and the ODV collection with the obtained climatic field. These tools can be useful for all who works with DIVAnd and wants to visualize results in ODV: one just needs to process .nc files with DIVAnd2ODV, which will create respective .csv files, then import the .csv file with first parameter in a normal way, and second – with "Merge" option.
- **DIVAnd update** to make the time varying background easier to use. An example of how to use it has been provided by ULiege (jupyter-notebook). Essentially: 1. Define the time selector

(month-list, year-list) for the background; 2. Make the background analysis; 3. Define the time selector (month-list, year-list) for the analysis; 4. Make the analysis with diva3d and the additional argument. Info on how to upgrade DIVAnd has been added as well in the software documentation.

- **Save observations to climatology file:** it raises the problem of **mapping metadata from WOD and CORA to SDN standards**. LOCal CDI ID and EDMO code are needed for backtracking. What about WOD and CORA observations which do not have CDI ID and EDMO code? Coriolis EDMO code (4630) has been assigned to CORA (MED and BLS products), but the other climatologies contain empty fields. A. Barth proposed to use EDMO code 1977 for WOD, corresponding to the NOAA/WDC) and the WOD identifier as local CDI (this was implemented in the BLS product). In CORA, the data have unique identifier too which is kept between releases and it could be used.
- **Generate XML file for SEXTANT:** DIVAnd *divadoxml* function can be used to generate product description for Sextant catalogue. It takes information from NetCDF file(s) with climatology and generates XML file according to a specific template (SeaDataNet template in this case). The XML file can then be imported directly to Sextant catalogue and presented as product description. In this release of climatology *divadoxml* was tested but was not used for Sextant pages. The main reasons for this were not fixed yet issue of WOD/CORA --> SDN metadata mapping, and missing information about data from WOD and CORA in the SeaDataNet CDI list. However in the next release of the climatology the possibilities of this function can be utilized in full scale as the workaround regarding metadata mapping have been proposed (see paragraph above), while the SeaDataNet CDI list can be manually extended with two records for WOD and CORA datasets.

  In the current release of climatology the product pages in Sextant were produced manually in the following way: the regional product page from the previous (SeaDataNet) climatology release was duplicated and used as a template, the data originators list was replaced with list from SeaDataCloud Data product, while the rest of the information was edited/added manually including the two new data originators of WOD and CORA datasets (NOAA/WDC and Coriolis respectively).
- **Use of misfits for a posteriori QC of observations:** in MED and BLS domains a two phases analysis has been implemented. The first analysis is needed to compute the residual observations. The residual outliers indices are then removed from the observations data set, the input file for the second phase analysis. The removal of observations with large residuals improves the quality of the final climatology. This iterative process has been tested and successfully applied .

3.1. Background field definition

Hereafter are reported the answer to the question: Sliding-window background estimate or rather a long-time average background and thus time invariant background field?

- Ö. Bäck (SMHI), BAL regional leader suggested to use a single background field from a long time average to be able to get as much data for it as possible, so a fixed time 1955-2017 for the background field. The same approach was used for BLS region climatology.
- H. Sagen (IMR) , ARC regional leader suggested use 30 year background field, 1955-1984, 1985-2014 instead of the 60 year background field. Then use those two 30 year for the 3 fixed decades.

- C. Troupin (ULiege): the choice for the period for **the background field may depend on the data coverage**: with the background (1) you want to get **the best "first-guess" of the analysi**s and at the same time (2) you want to have observations distributed over all the domain. With a long period (60 years for example), the background will certainly satisfy the second condition and have data almost everywhere. With a shorted period you can expect to have better first-guess, for example the 1955-1984 is probably better than the 1955-2014 if you work on 1955-1965, but if there are many regions without data in the background, then it is better to switch to a longer period. (3) Another case to consider is the Mediterranean, where the conditions changed quite a lot over a relatively short period of time (Eastern Mediterranean Transient), in that case it could also makes sense to have one background representative of the conditions before the Transient, and one after. It would be good to have some harmonization between the product.
- A. Barth (ULiege): in the Arctic are indeed quite significant since the ~1990 and it would make sense not to couple these decades by the background (at least for the Arctic). Indeed the Mediterranean is also another well documented case with a long-term trend. **The approach of two 30-year background fields: 1955-1984, 1985-2014 is suitable for other domains (eventually with different year ranges).**

3.2. Vertical stability

When merging the Temperature and Salinity fields the resulting Density profiles might have vertical instability in upper layer. Density inversion is a tricky issue, a post-processing tool could be developed which would diffuse vertically unstable profiles. The "stabilisation" procedure implemented in the old Diva was tested in specific cases (it is not always easy to find those situations, as we wanted to find locations where the unstable profiles were created by the interpolation, but not were not initially present). Then the code evolved but that procedure was never applied in production for the climatologies. The doc can be accessed at

http://modb.oce.ulg.ac.be/mediawiki/upload/DIVA/notes/DivaSTBUserGuide.pdf

Modifying the *vertical correlation length* (VCL) either in the background field and the analysis could be a possible alternative solution too but it is not straightforward. In fact, the tool which estimates the vertical correlation length based on the vertical density gradient from a preliminary gridded analysis could be an option. It needs to redo the analysis with this adjusted correlation length. But this process might be quite slow as it involves running DIVAnd multiple times, thus the "vertical diffusion approach" is preferable.

In DIVA, there was the possibility to add pseudo observations to avoid inversion. However, this code is no longer part of DIVAnd (or at least maintained). This option will be evaluated for future changes.

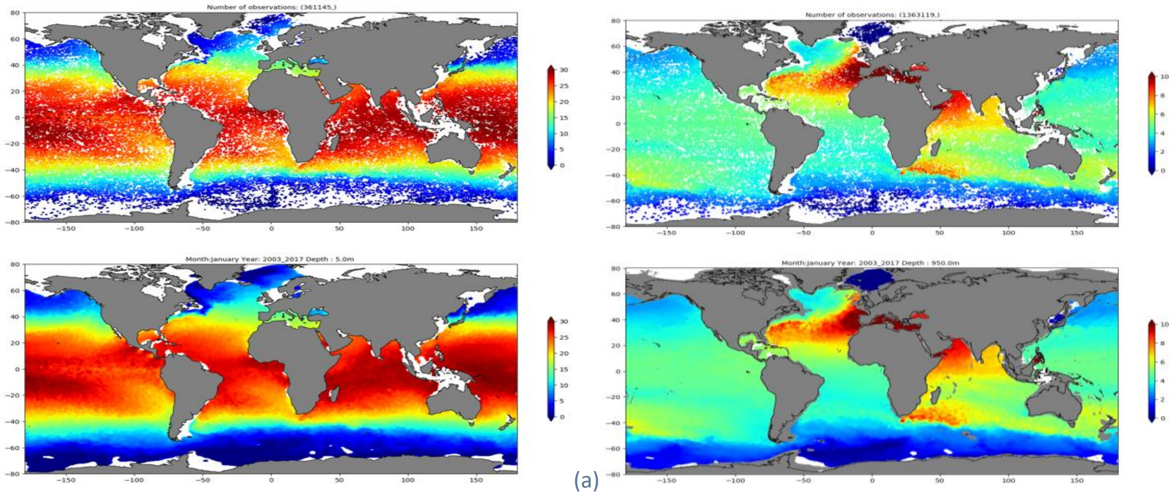
The WP team decided to do not apply any post processing tool at this stage and release the original fields. A good option in the future would be to release both or the original TS and the corrected fields or the tool to adjust it.

4. Results

The results are presented per sea region with some plots extracted from each PIDoc, as examples. Please refer to the table in Appendix 1 for products' characteristics.

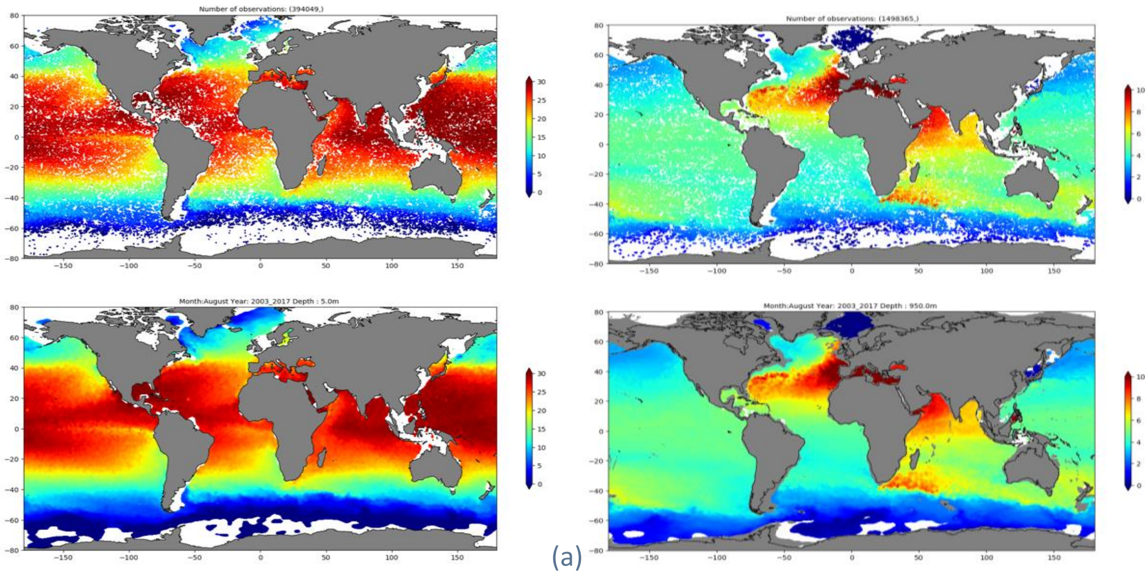
4.1. Global Ocean

January and August Temperature climatology for (SDC_GLO_CLIM_T_V1_2) at the surface and 950 m are displayed in Figure 3 and Figure 4.



(a) (b)

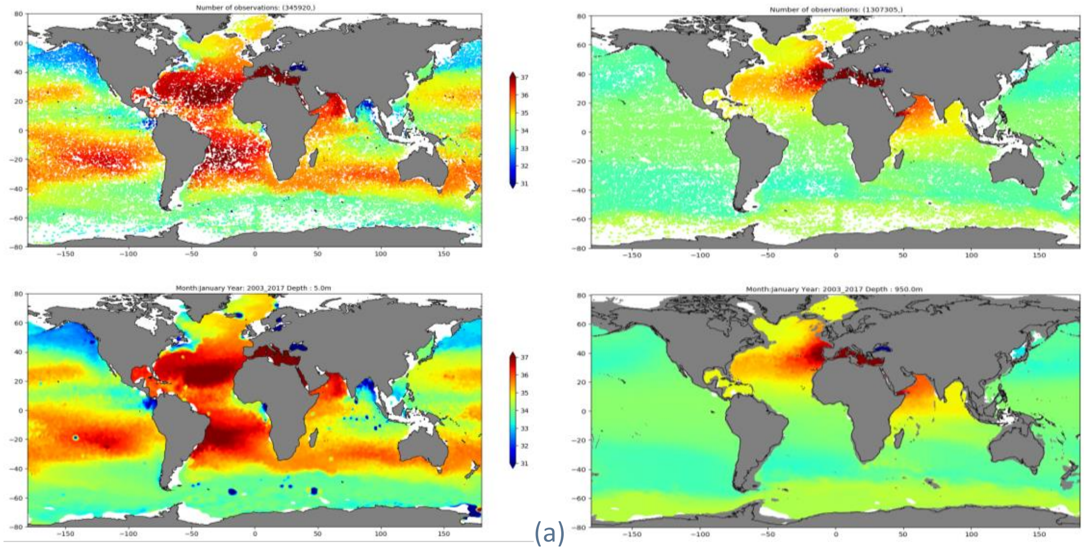
Figure 3 - SDC_GLO_CLIM_T_V1_2 January Temperature field in Degree C: a) at surface; b) at 950 m.



(a) (b)

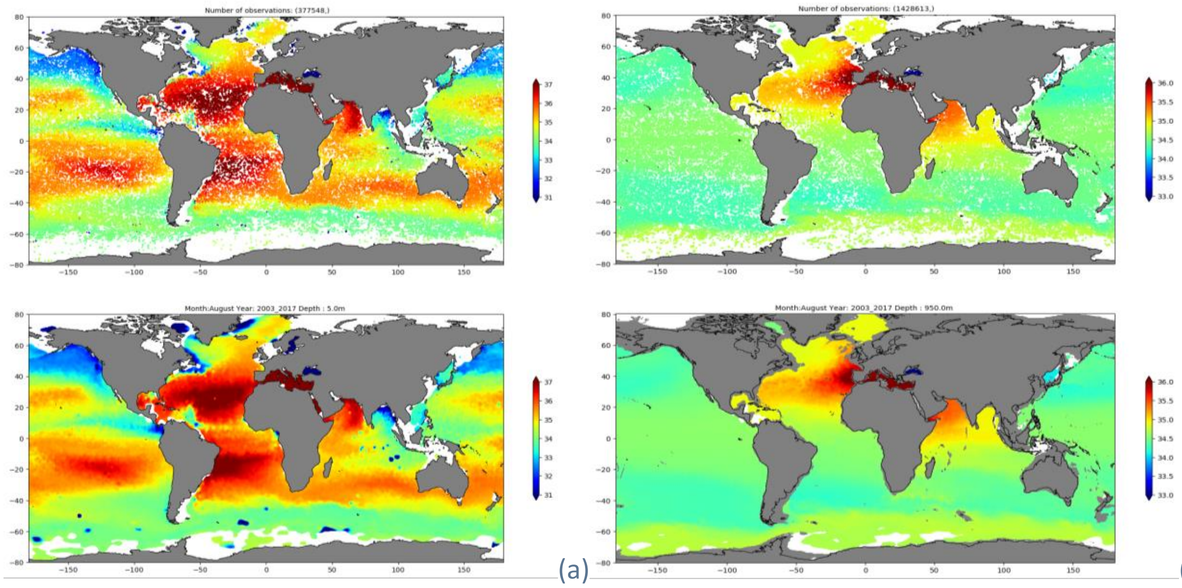
Figure 4 - SDC_GLO_CLIM_T_V1_2 August Temperature field in Degree C: a) at surface; b) at 950 m.

Salinity climatology SDC_GLO_CLIM_S_V1_2 is displayed in Figure 5 (January) and Figure 6 (August).



(a) (b)

Figure 5 - SDC_GLO_CLIM_V1_2 January Salinity field in PSU: a) at the surface; b) at 950m.

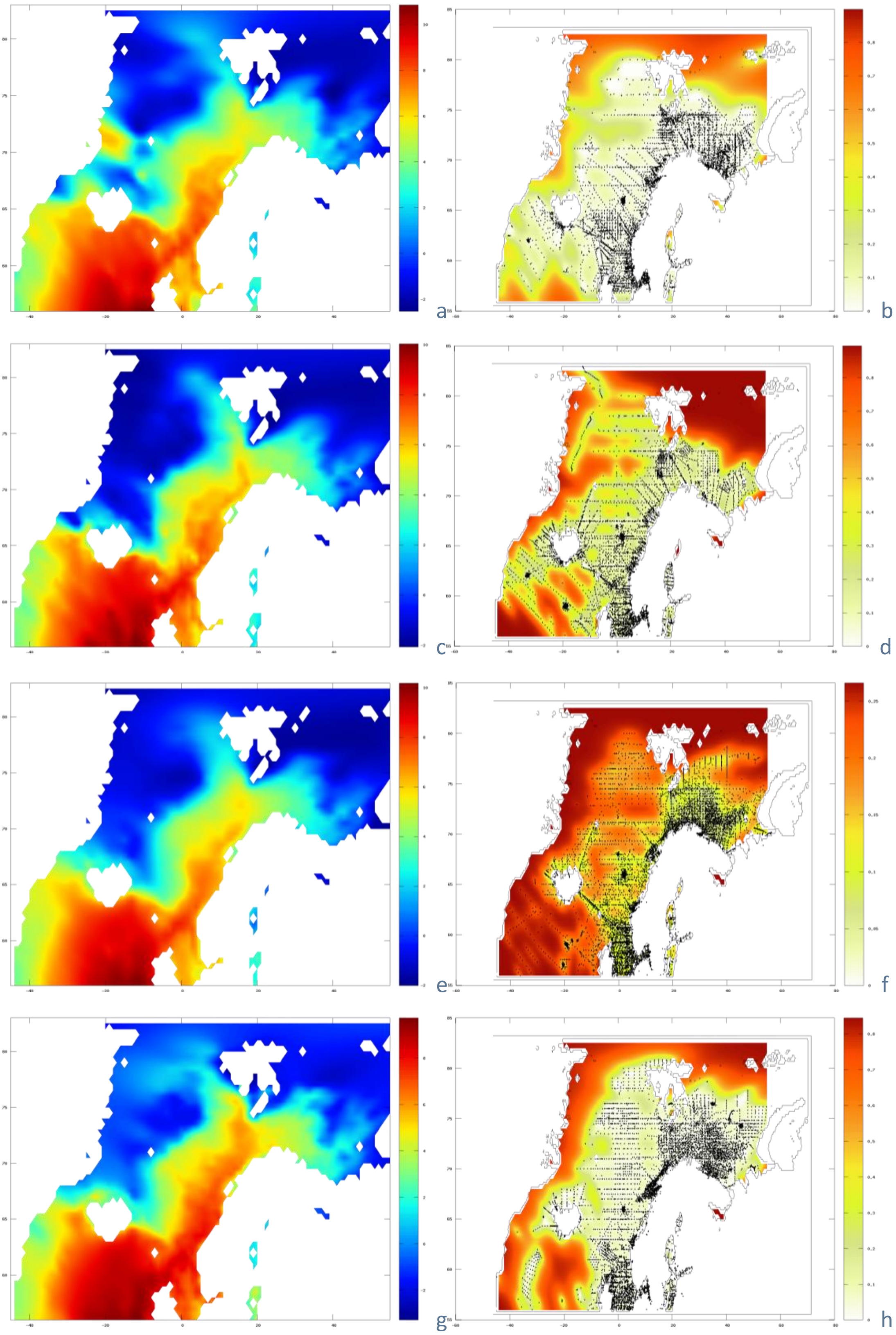


(a) (b)

Figure 6- SDC_GLO_CLIM_V1_2 August Salinity field in PSU: a) at the surface; b) at 950m.

4.2. Arctic Ocean

Figure 7 shows the ARC winter climatology at 50m depth the relative error field as a function of observations' location.


a
b
c
d
e
f
g
h

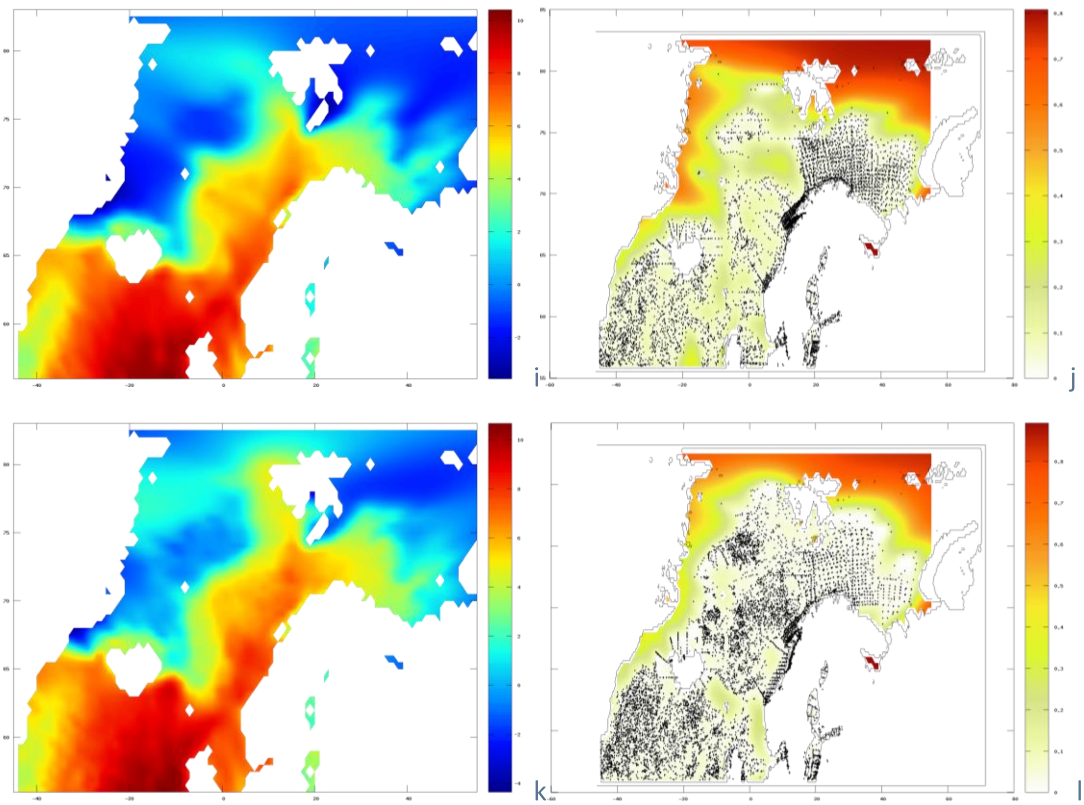


Figure 7 - SDC_ARC_CLIM_T_V1_1 winter climatology for temperature at depth 50m for the 6 decades (from top to bottom 1955-1964, 1965-1974, 1975-1984, 1985-1994, 1995-2004, 2005-2014): (a-c-e-g-I-k) decadal Temperature fields at the surface without error masking; (b-d-f-h-j-l)) Error field and observations' points.

4.3. North Sea

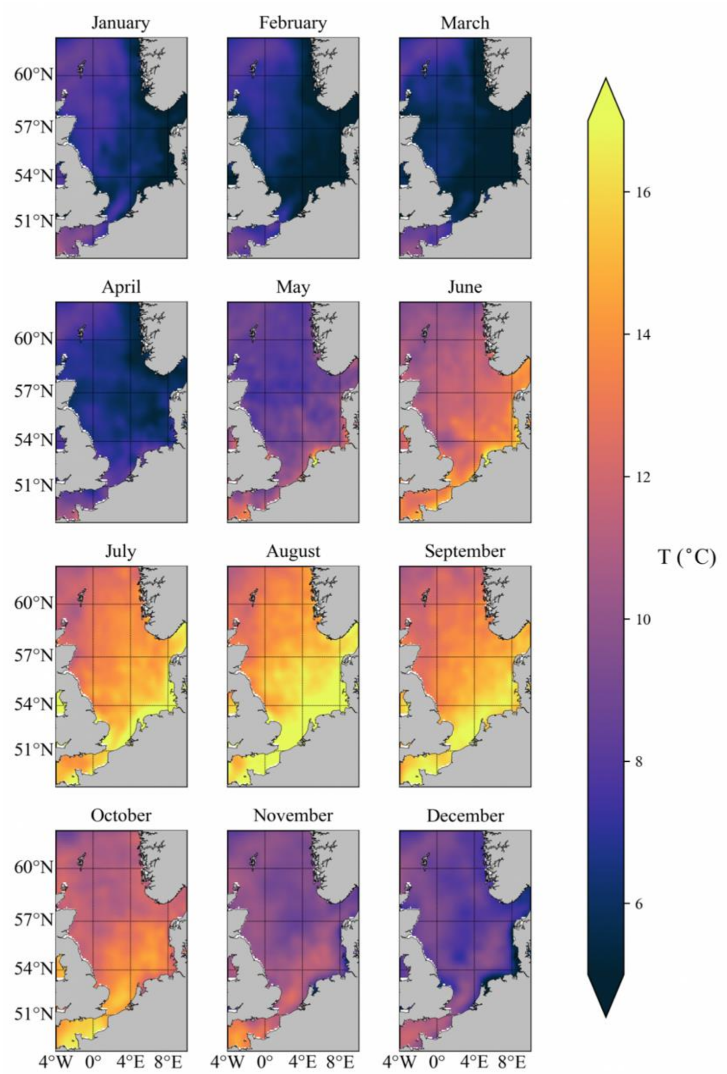


Figure 8 - SDC_NS_CLIM_T_V1_1 monthly Temperature fields at the surface as computed with the whole data set (1955-2014).

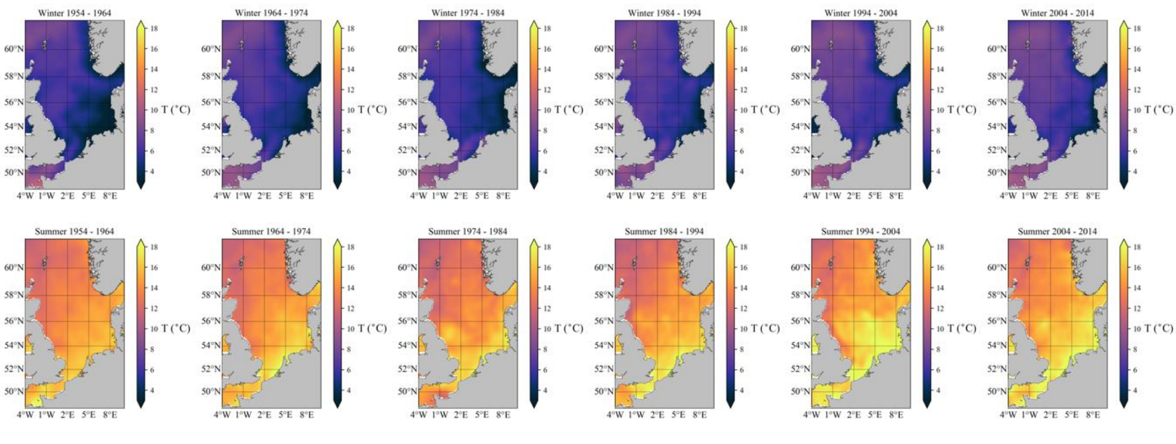


Figure 9 - SDC_NS_CLIM_T_V1_2 Variation of winter Temperature (Jan, Feb, Mar-top) and summer Temperature (Jul, Aug, Sep-bottom) at the surface in 6 decades (left to right 1955-1964, 1965-1974, 1975-1984, 1985-1994, 1995-2004, 2005-2014).

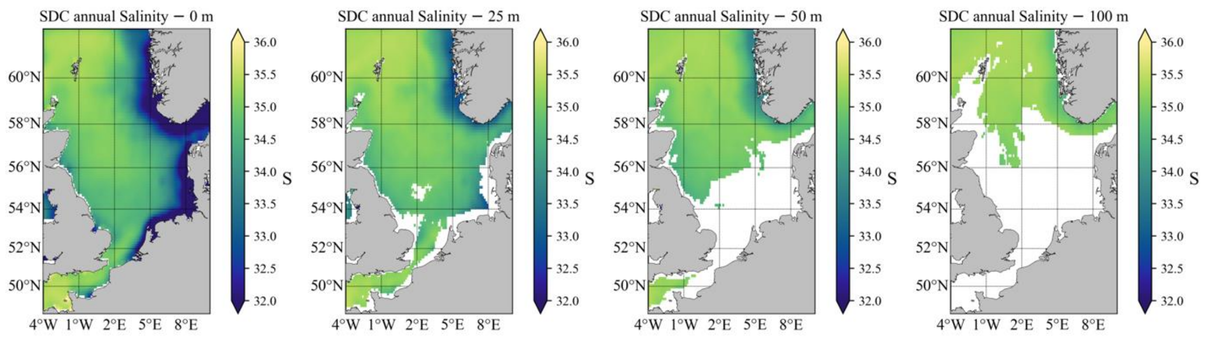
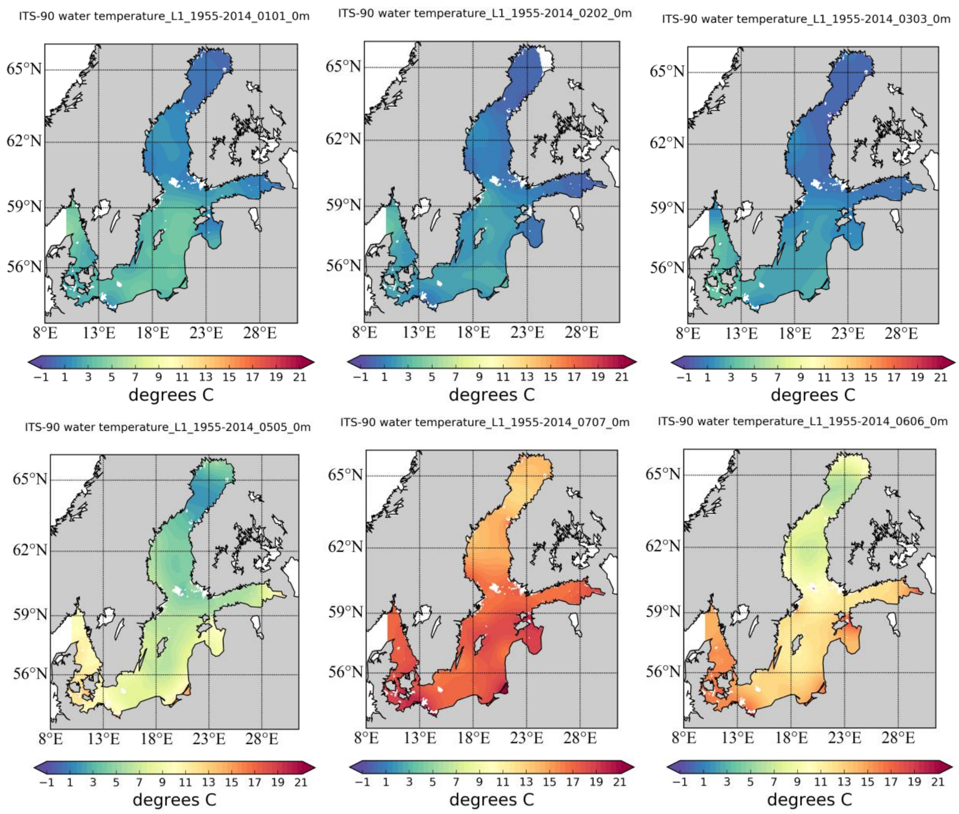


Figure 10 - SDC_NS_CLIM_T_V1_1 annual salinity fields at depths = 0, 25, 50 and 100m as computed with the whole data set (1955-2014).

4.4. Baltic Sea

Figure 11 shows the monthly climatological fields of temperature at the surface, from January, upper left, to December, lower right. Temperature varies from below zero in the northern region, during late winter and early spring, to about 20 °C in the southern region during summer. The warmest temperatures occur in August.

Figure 12 displays August decadal variation of Salinity in the Baltic Sea.



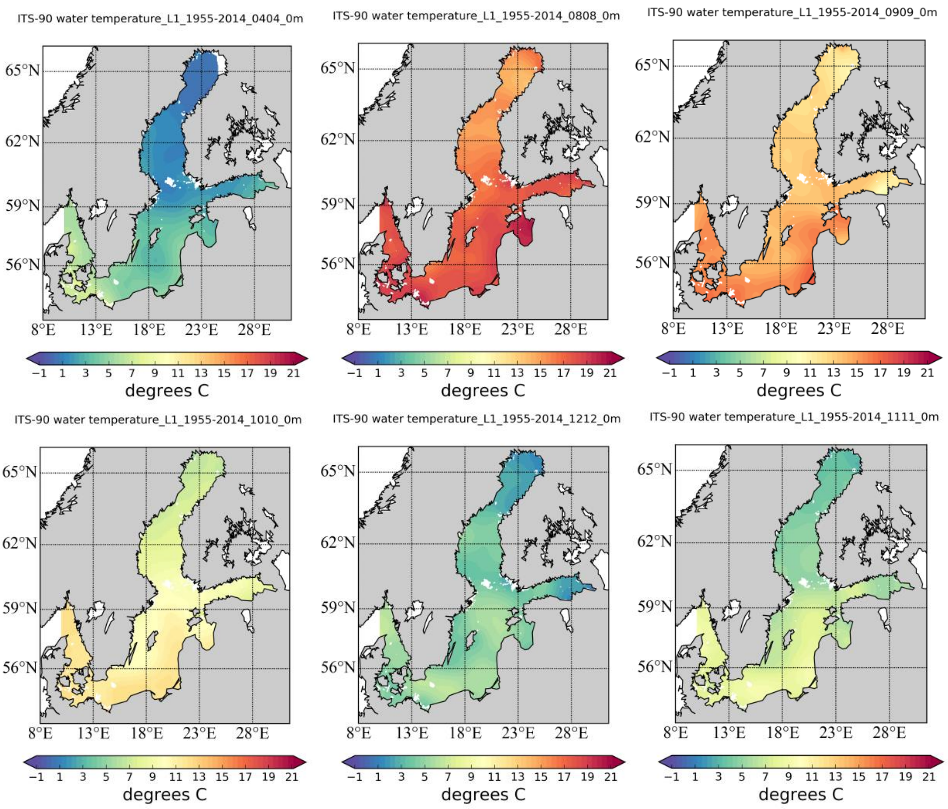


Figure 11 - SDC_BAL_CLIM_TS_V1_2 monthly temperature at surface for 1955-2014 time period, from January, top left, to December, bottom right.

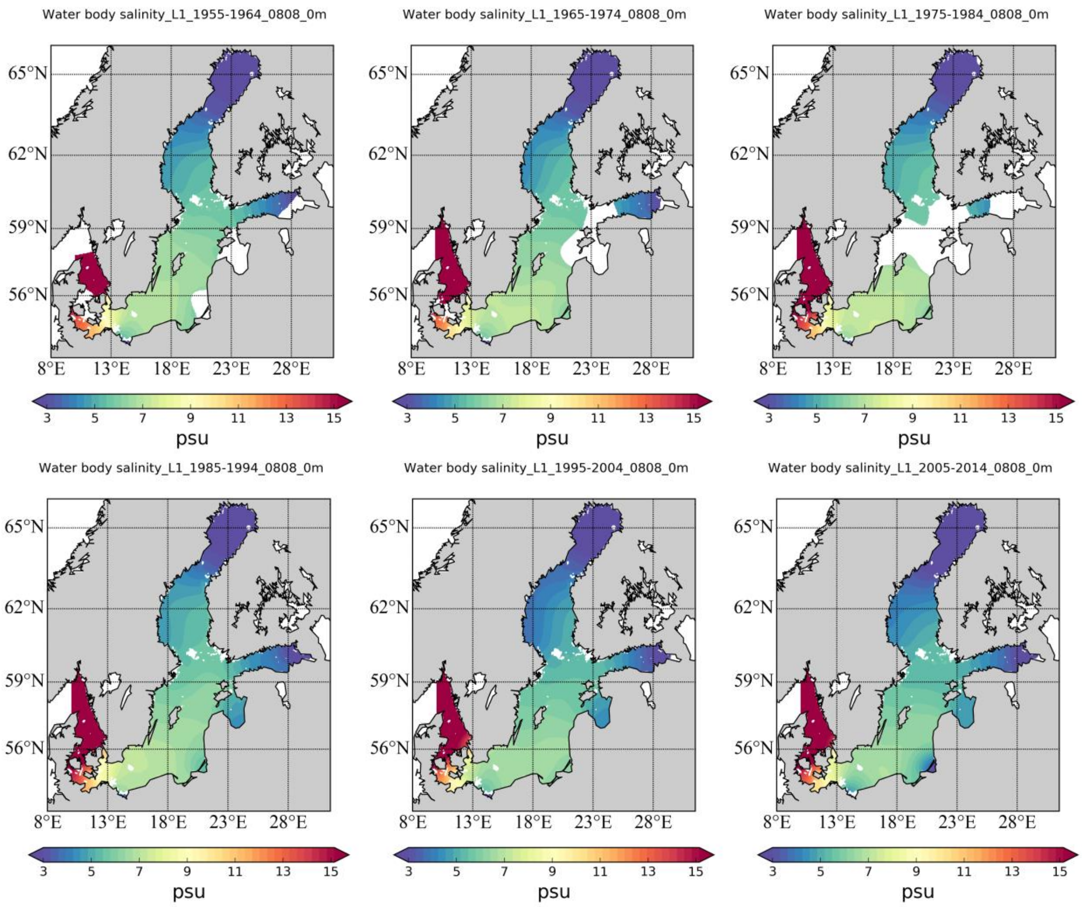


Figure 12 - SDC_BAL_CLIM_TS_V1_1 Surface salinity in August per decade from 1955-1964 (top left) to 2005-2014 (bottom right).

4.5. North Atlantic

Fields at two vertical levels have been chosen at the surface and at 1000m of depth, to show respectively the resulting surface T and S distributions and the circulation of the Mediterranean Outflow Waters in the North Atlantic.

Temperature fields in January and July show differences; warmer surface currents and cold North Atlantic Deep Water (NADW) can be easily identified. The thermohaline circulation heats the North Atlantic and Northern Europe. It extends right up to the Greenland and Norwegian Seas, pushing back the winter sea ice margin (Figure 13).

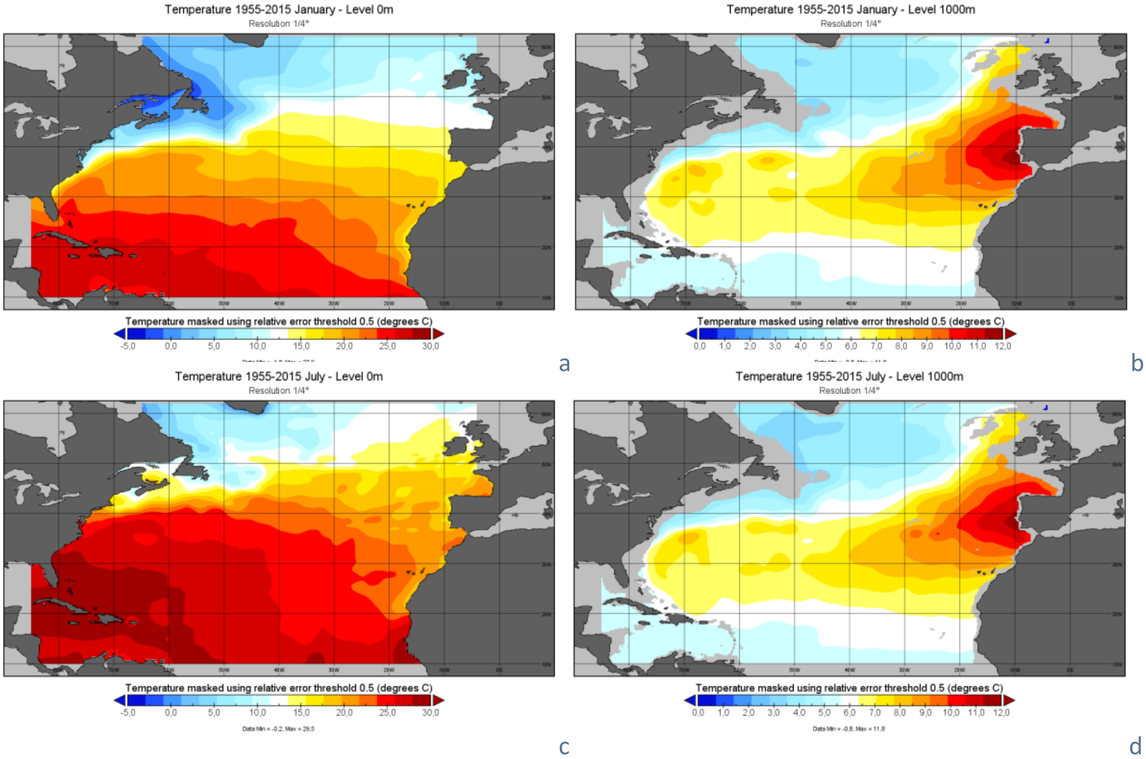


Figure 13 - Temperature fields: a) January at the surface; b) January at 1000m; c) July at the surface; c) July at 1000m depth.

Figure 14 displays the decadal fields at seasonal resolution (seasons being defined as: Winter 1202 Spring 0305, Summer 0608, Autumn 0911) from decade 1955-1964 to decade 2005-2015, the temperature structure becomes finer due to the increasing data coverage with time.

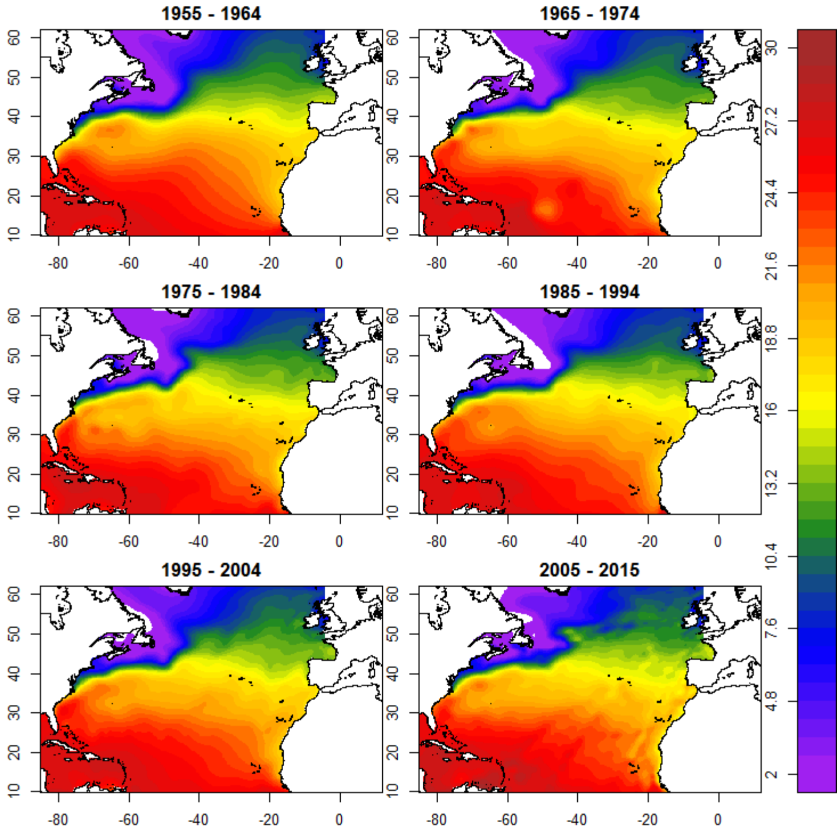


Figure 14 - Temperature map at the surface for the 6 decades in Spring (from March to May).

The distribution of the salinity fields at two depth levels have been chosen in Figure 15 to focus on the surface patterns and at 1000m where the circulation of the Mediterranean Outflow Waters can be tracked in the Atlantic Ocean due to a salinity maximum. The decrease of salinity in the North Atlantic Ocean is concentrated in the west and linked with advection by the East and West Greenland Currents and the Labrador Current. Subsurface salinity minima (S<34.95) originating from the Labrador Sea Water (LSW) core are observed on the American continental slope. These minima are traces of the LSW flowing southward along the upper bound of the NADW water mass (Figure 15). LSW is characterised by a minimum of salinity and this signature is found in the western and eastern subpolar North Atlantic.

Both salinity and temperature maxima are observed at 1000 m depth near the upper distribution limit of NADW, this is the sign of the Mediterranean Outflow Water, which is carried northward along the Portuguese shelf and mixes into the subtropical gyre circulation.

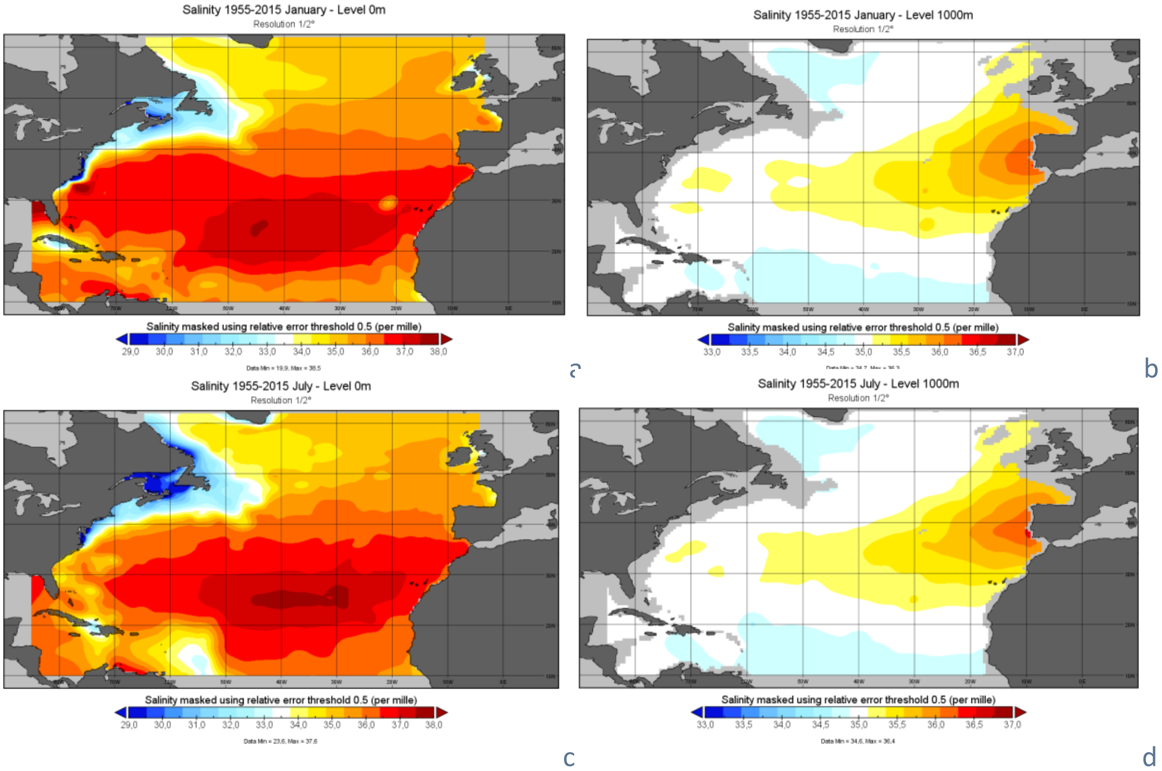


Figure 15 - Salinity fields: a) January at the surface; b) January at 1000m; c) July at the surface; c) July at 1000m depth.

Decadal salinity at seasonal resolution in winter Figure 16 show the spatial structures become finer due to the increasing data coverage with time.

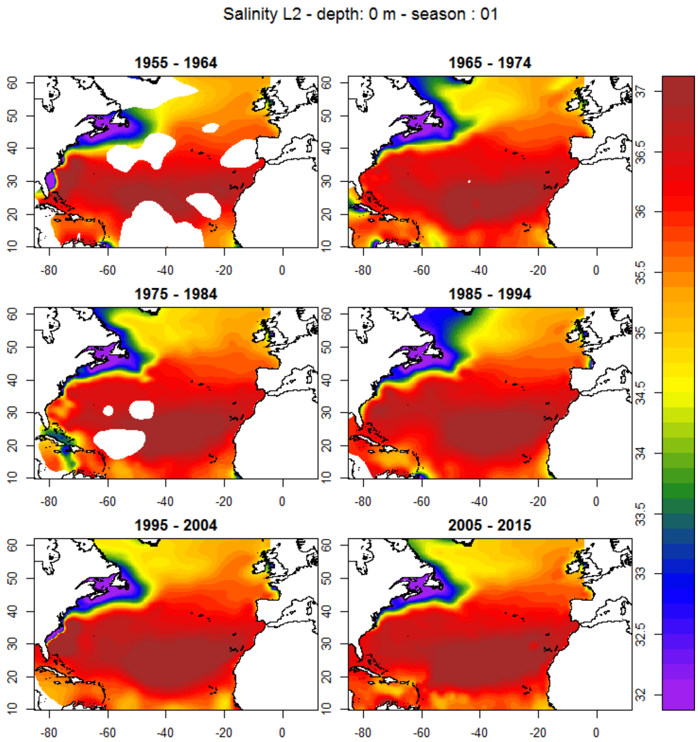


Figure 16 - Winter salinity maps at the surface for the 6 decades (from December to February).

4.6. Mediterranean Sea

The Salinity monthly climatological fields at 300m of depth, representative of the intermediate waters in the Mediterranean Sea, are shown in Figure 17, Figure 18 and Figure 19 for three time period respectively:

- 1955-2017 SDC_MED_CLIM_TS_V1_1
- 1955-1984 SDC_MED_CLIM_TS_V1_2: pre Eastern Mediterranean Transient
- 1985-2017 SDC_MED_CLIM_TS_V1_3: post Eastern Mediterranean Transient.

Comparing Figure 18 and Figure 19 it is apparent the increase of salinity in the time period 1985-2017, when EMT (early nineties) determined the formation of saltier intermediate and deep water masses than before, especially in the Eastern Mediterranean basin.

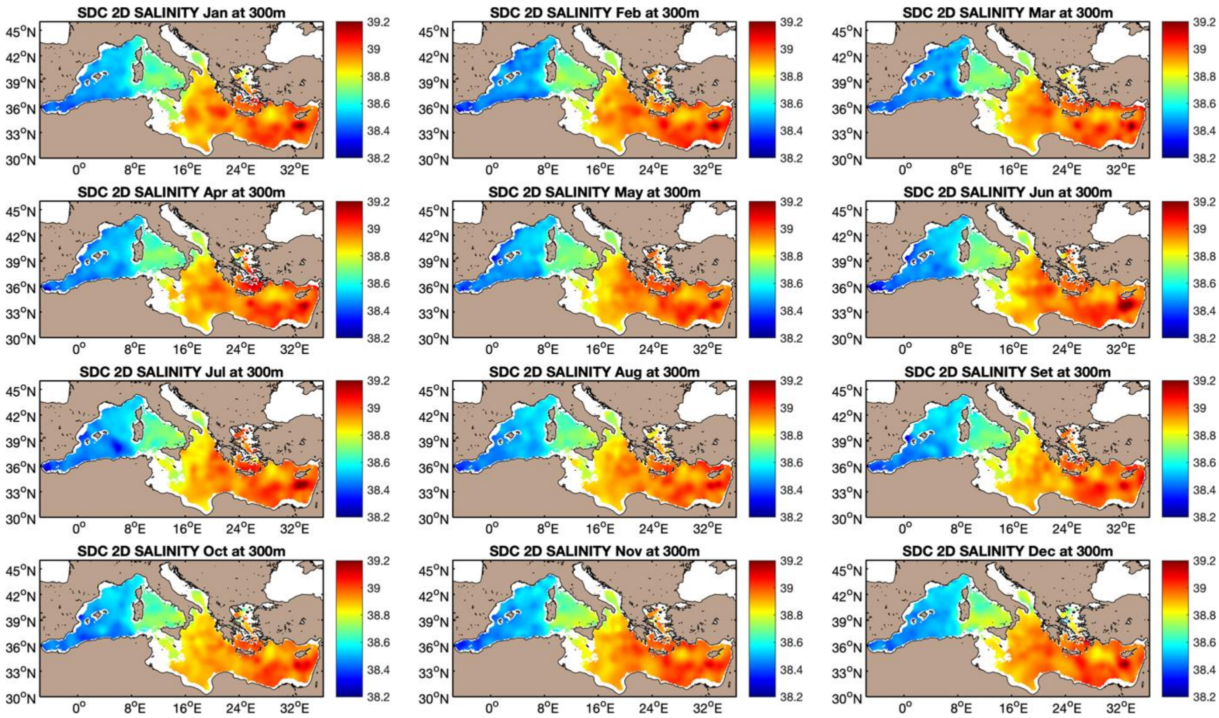


Figure 17 - SDC_MED_CLIM_TS_V1_2 monthly salinity at 300m (intermediate water) for 1955-2017 time period from January (top left) to December (bottom right).

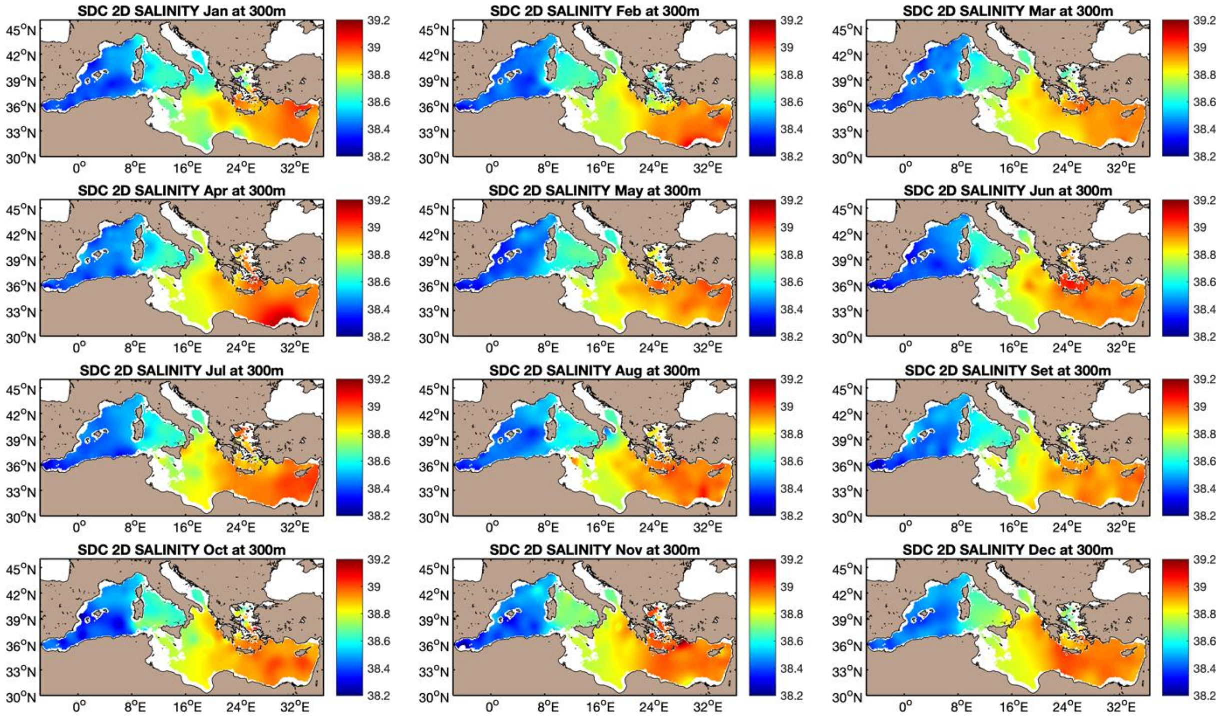


Figure 18 - SDC_MED_CLIM_TS_V1_2 monthly temperature at 300m (intermediate water) for 1955-1984 time period (pre-Eastern Mediterranean Transient), from January (top left) to December (bottom right).

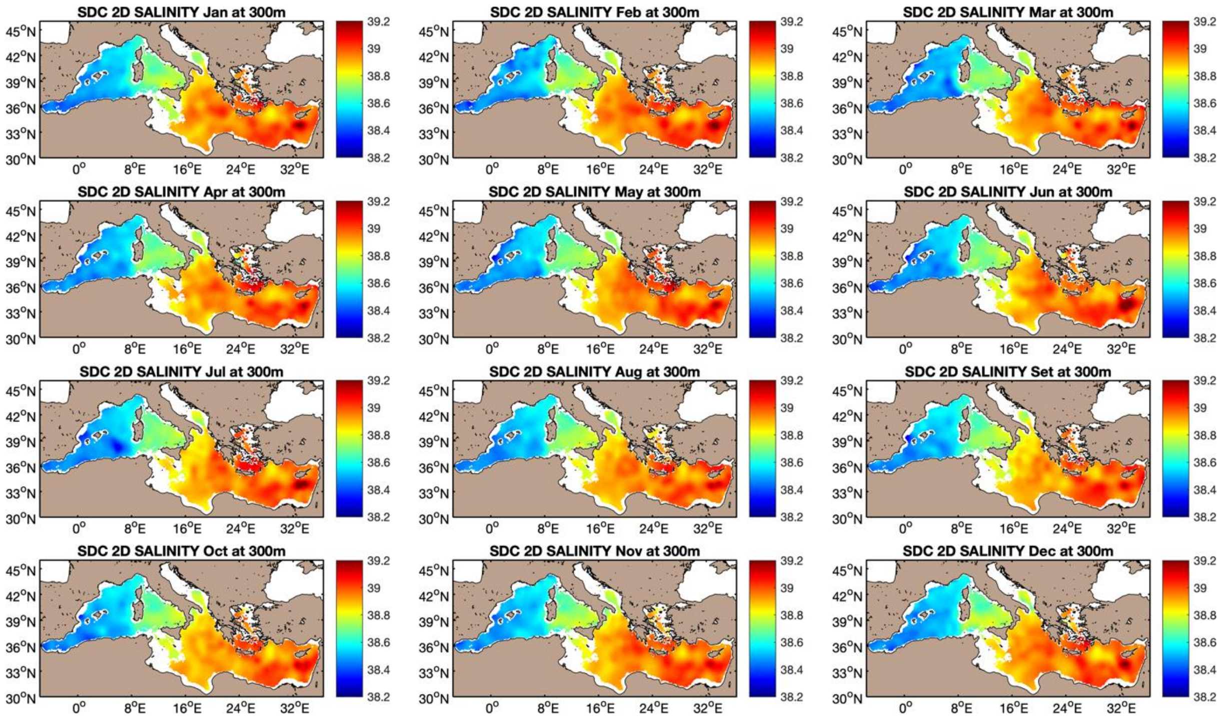


Figure 19 - SDC_MED_CLIM_TS_V1_3 monthly temperature at 300m (intermediate water) for 1985-2017 time period (post-Eastern Mediterranean Transient), from January (top left) to December (bottom right).

4.7. Black Sea

Figure 20 shows the monthly climatological fields of temperature at the surface: winter temperatures vary from -1 C° in the NW part and in the Sea of Azov to 10 C° in South. The summer temperatures reach 28 C°.

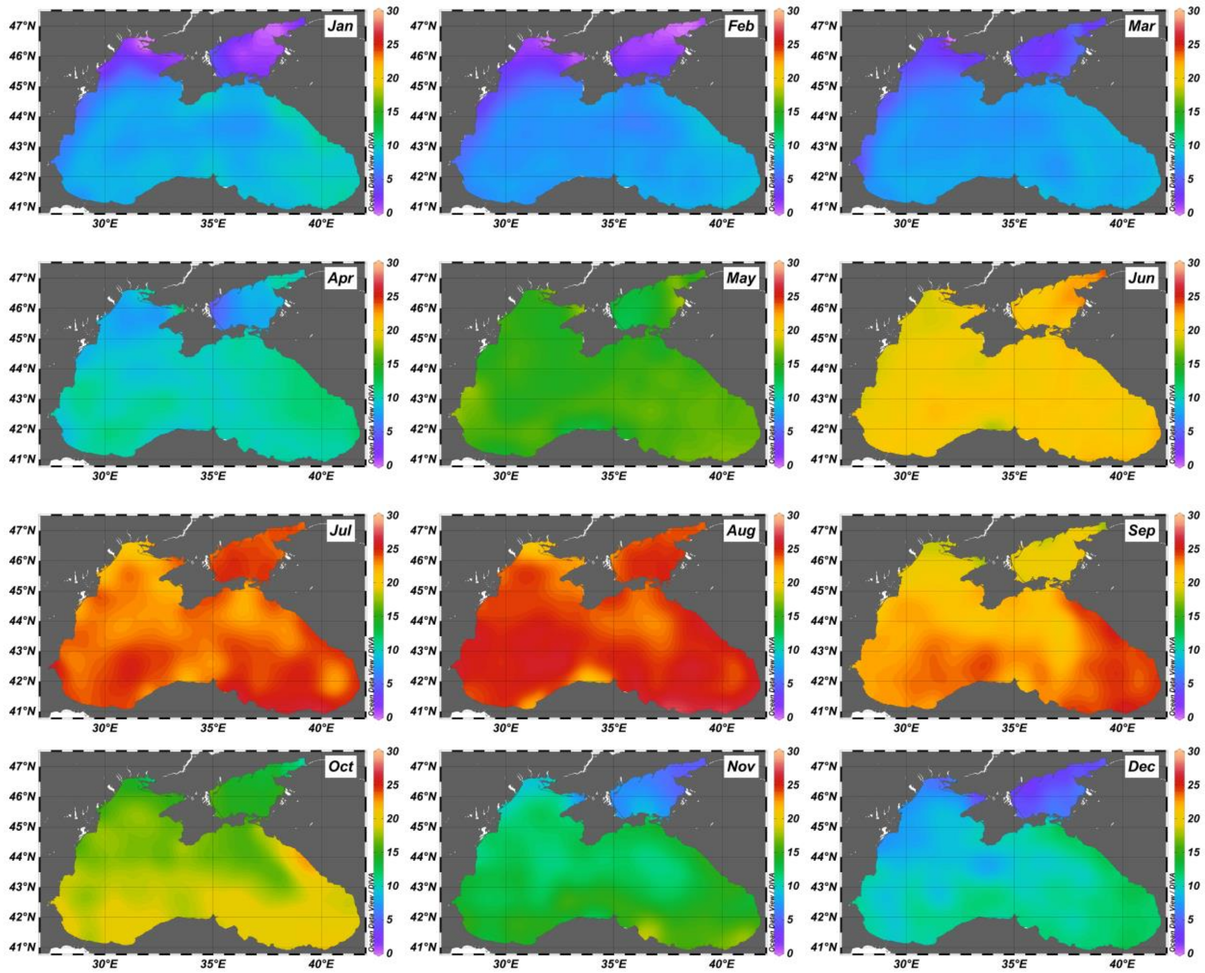


Figure 20 - Monthly variation of Temperature at surface (DIVAnd analysis for 1955-2017).

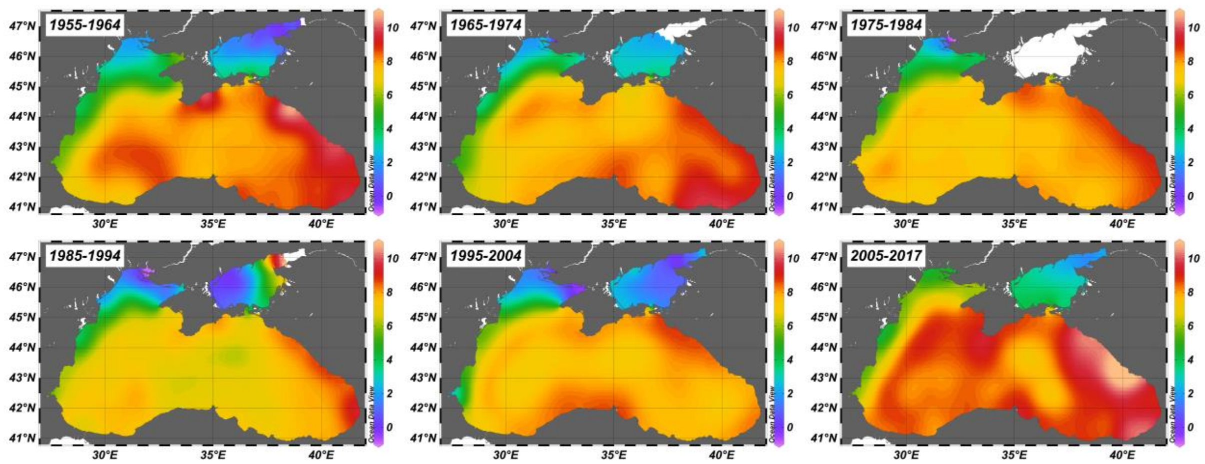


Figure 21 - Variation of winter Temperature (Jan, Feb, Mar) at the surface in 6 decades over the time period 1955-2017.

The shallow NW part of the Black Sea and the Sea of Azov receive most of the river fresh water inflow, and salinity here goes down to 11 and up to 0 in river estuaries. The annual cycle of fresh water input is well traced in Figure 22.

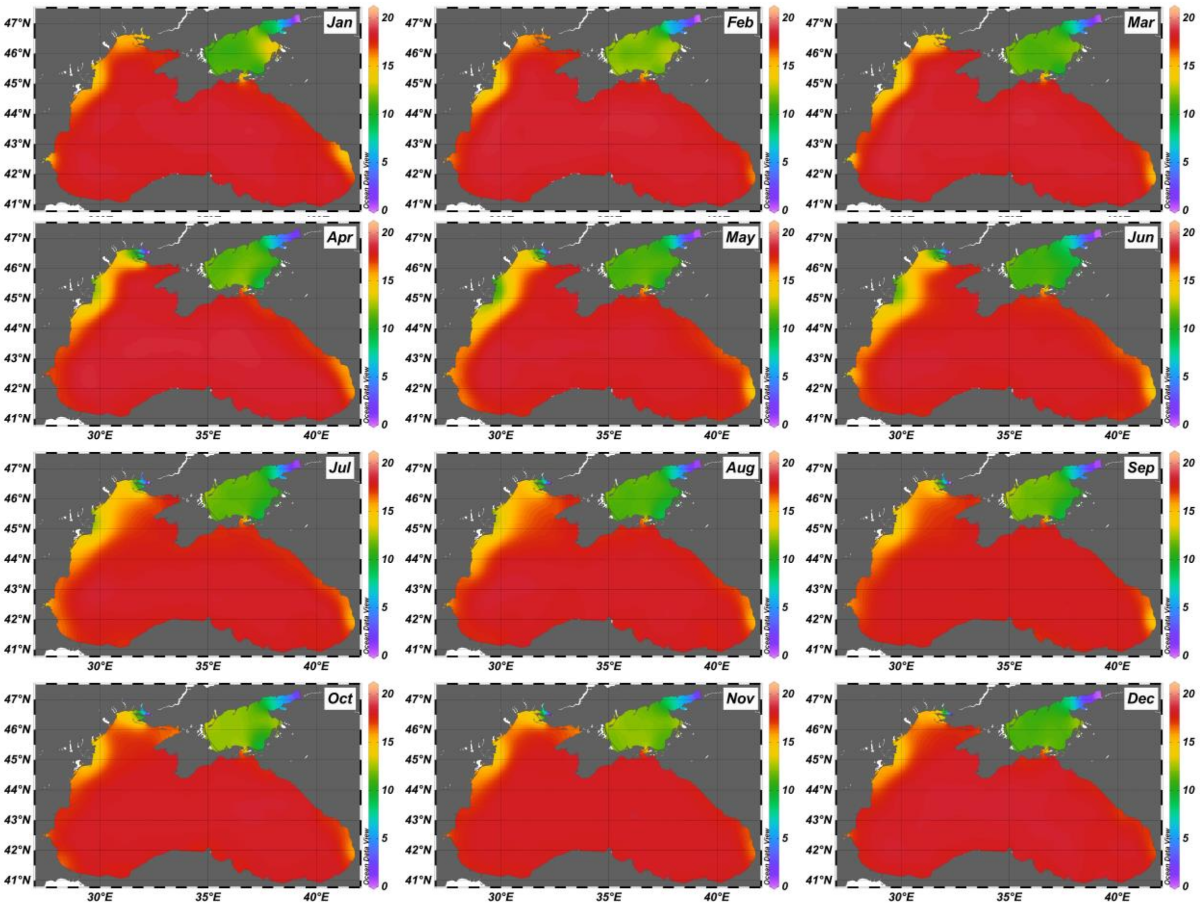


Figure 22 - Annual variation of Salinity at the surface (DIVAnd analysis for 1955-2017)

5. Consistency Analysis

The validation of the analysis is an essential step in order to have an indication of the reliability of the results. For the validation/consistency analysis we decided to adopt the methodology conducted in the Mediterranean region in the framework of SeaDataNet2 (see D11.5 for details). We considered as a reference data WOA. GLO product has been compared with WOA2013 V2 (Locarnini et al., 2013 and Zweng et al., 2013), while regional products have been compared to WOA2018 (Locarnini et al., 2018 and Zweng et al. 2018). A motivation for GLO product choice is the usage of WOD 2013 as input data set for the climatology generation.

A qualitative consistency analysis has been performed through visual check of images for ARC, BAL, NS, NAT products, while GLO, MED and BLS products were also compared quantitatively through indices. The quality indices calculated are root mean square difference (RMSD), Bias and for GLO and MED product also the percentage root mean square difference (PRMSD) computed as follow:

$$RMSD = \sqrt{\frac{1}{N}\sum(x-y)^2}$$

$$BIAS = \frac{1}{N}\sum(x-y)$$

where x is the WOA field and y is the SDC field, and

$$PRMSD = \frac{RMSD}{REF}$$

where

$$REF = \sqrt{\frac{1}{N}\sum(x - \bar{x})^2}$$

is the standard deviation of the WOA monthly climatology, N are the number of points of the SDC analysis grid and $\bar{x}$ is the average.

Both qualitative and quantitative analysis suggested a good consistency of SDC products with WOA and highlighted the added value of regional products compared to the global WOA at lowest spatial resolution.

Hereafter are reported two examples of consistency analysis for the global ocean and the Black Sea.

5.1. Consistency analysis SDC_GLO_CLIM_TS_V1 climatology

In the GLO case, the average of objectively analysed WOA13 V2 climatology at 0.25 for six decades (1955-1964, 1965-1974, 1975-1984, 1985-1994, 1995-2004 and 2005-2012) has been used. To compute the difference between the SDC_GLO_CLIM_TS_V1 climatology, WOA fields were interpolated on the SDC analysis grid using linear interpolation (python library). The differences between the interpolated WOA and DIVA, for temperature and salinity are shown in Figure 23 and Figure 24 respectively.

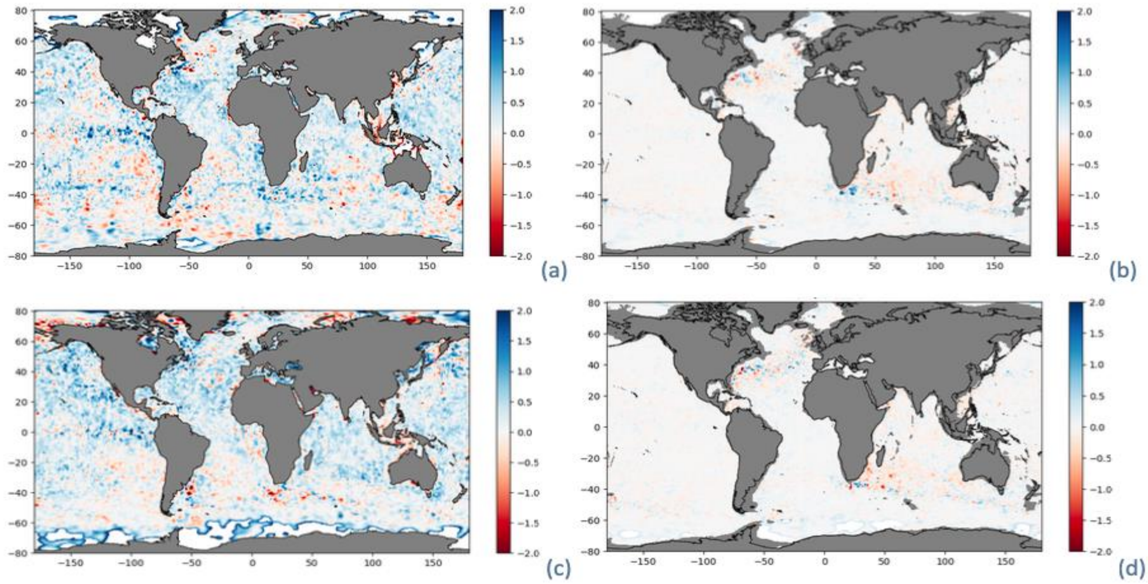


Figure 23 - Difference field (WOA-SDC_GLO_CLIM_V1_1) for Temperature: a) January at the surface; b) January at 950m; c) August at the surface, and d) August at 950m.

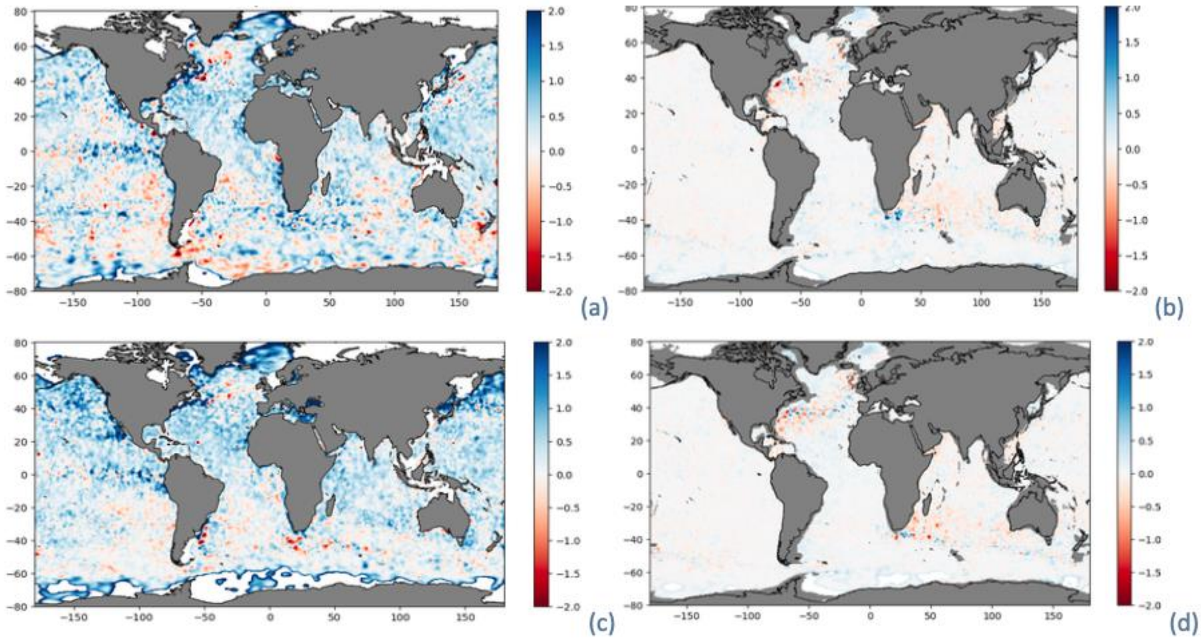


Figure 24 - Difference field (WOA-SDC_GLO_CLIM_V1_2) for Temperature: a) January at the surface; b) January at 950m; c) August at the surface, and d) August at 950m.

The GLO RMSD and Bias were calculated for surface, 950m and 1500m for all months.

The RMSD for SDC_GLO_CLIM_T_V1_1 varies from 0.43 to 0.51ºC for the surface layer, 0.13 to 0.14ºC for the 950m layer and 0.05 to 0.07ºC for the 1500m layer. Similarly, the Bias for surface layer ranges from -0.14 to -0.21ºC, -0.01 to 0.003ºC for the 950 m layer and -0.01ºC for the layer at 1500m. The PRMSD at the surface is about 5-6%, at 950m is 8-9% and at 1500m is 6-7%.

The RMSD for SDC_GLO_CLIM_T_V1_2 varies from 0.6 to 0.7ºC for the surface layer, 0.1 to 0.12ºC for 950m layer and 0.06 to 0.07ºC for 1500m layer. Similarly, the Bias for surface layer ranges from -0.27 to -0.4ºC, -0.02ºC at 950 m layer and -0.01ºC at 1500m layer. The PRMSD value at the surface is about 7-8%, at 950m is 8-9% and at 1500m is 6- 8%.

The RMSD for SDC_GLO_CLIM_S_V1_2 varies from 0.38 to 0.67 PSU for the surface layer, varies from 0.0012 to 0.01 PSU at 950m layer and 0.008 to 0.009 PSU for 1500 layer. Similarly, the Bias for surface layer ranges from -0.27 to -0.4 PSU, -0.02 PSU at 950 m layer and -0.01 PSU at 1500m. The PRMSD value for surface is 22% to 34%, at 950m is 3% to 4% and at 1500m is 3%. The reason for greater value of PRMSD is anomalous values in data that results as sharp gradients in the gridded field

The RMSD for SDC_GLO_CLIM_S_V1_2 varies from 0.39 to 0.82 PSU for surface layer, varies from 0.013 to 0.01 PSU for layers at 950m and 0.008 to 0.009 PSU for 1500m layer. Similarly, Bias for surface layer ranges from 0.01 to 0.06, -2e-4 to 1.75 PSU for layer at 950 m and at is 2e-4 to 5e-4 PSU for 1500m layer. The PRMSD value for surface is 18% to 67%, at 950m is 3% to 9% and for layer at 1500m is 3% to 6%. The reason for greater value of PRMSD is anomalous values in data that results as sharp gradients in the gridded field.

5.2. Consistency analysis SDC_BLS_CLIM_TS_V1 climatology

Comparison of selected Temperature and Salinity fields is presented in Figure 25. The maps have similarities and differences. The WOA18 maps are smoother and have less detail, while SDC maps look noisier but seem to be more realistic. For example, in SDC temperature map the temperature gradient in NW Black Sea in January is stronger; in SDC salinity map the areas of river inflows are more pronounced, the Sea of Azov has a more natural salinity distribution.

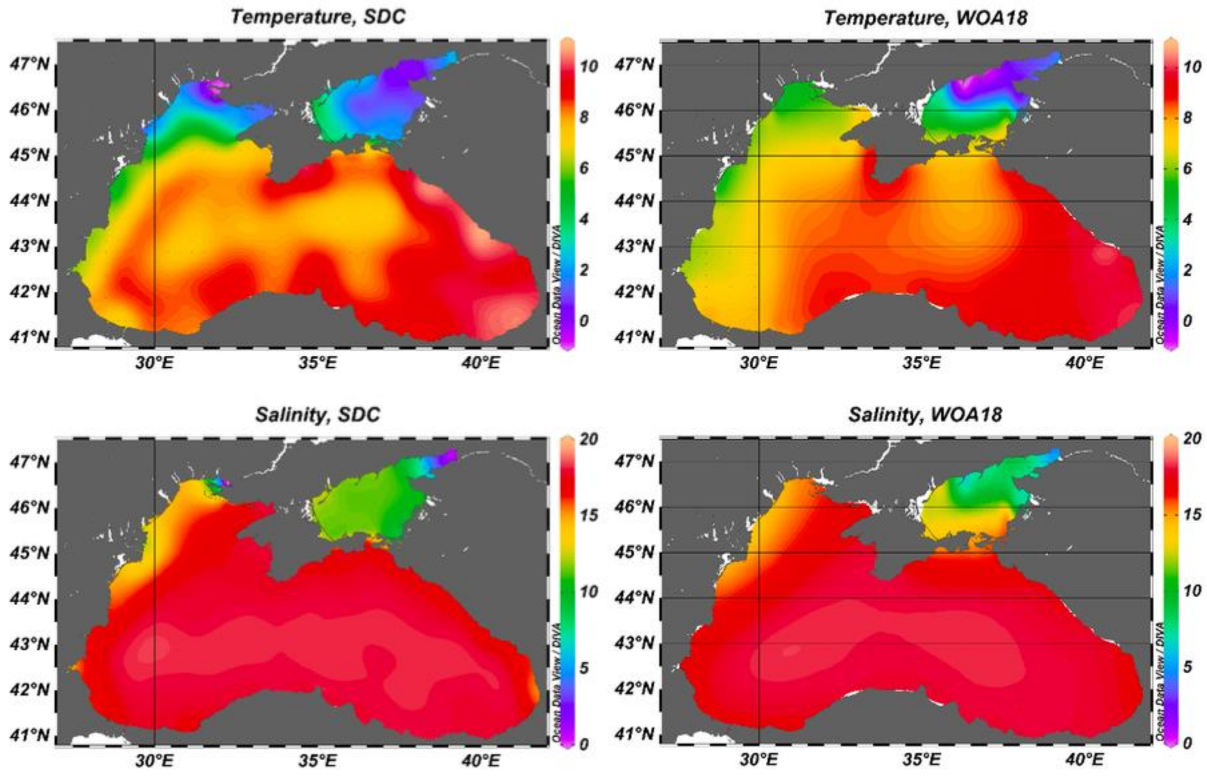


Figure 25 - SDC and WOA18 temperature (January) and salinity (July) maps at the surface for time span 1955-2017.

To quantify the differences between WOA18 and SDC climatologies the statistical indexes BIAS and RMSE were calculated for monthly fields for time span 1955 – 2017 and for seasonal fields – the decadal ones and for time span 1955 – 2017. Comparison was performed for matching grid nodes at 0.125, 03.75, 0.625, and 0.875 degree. The global BIAS indexes (i.e. calculated for the 3d grid and whole time span) in Table 4 are close to zero suggesting overall correspondence of WOA18 and SDC fields, however relatively large RMSE indexes suggest presence of significant differences, possibly dependent on depth and time.

More detailed analysis was performed for monthly fields. The main differences are observed in upper layer 0 – 300 m, while below 300 m both BIAS and RMSE are close to zero that explains low values in Table 4.

Table 4 - Statistical indexes of difference between WOA18 and SDC climatology

Fields	Time span	Temperature		Salinity	
		BIAS	RMSE	BIAS	RMSE
Seasonal	**1955 - 1964**	-0.07	0.54	0.01	0.21
Seasonal	**1965 - 1974**	-0.06	0.51	-0.01	0.21
Seasonal	**1975 - 1984**	0.08	0.50	-0.06	0.22
Seasonal	**1985 - 1994**	0.03	0.42	-0.03	0.21
Seasonal	**1995 - 2004**	-0.09	0.89	-0.08	0.33
Seasonal	**2005 - 2017**	-0.03	0.47	0.03	0.28
Seasonal	**1955 - 2017**	0.02	0.39	-0.03	0.22
Monthly	**1955 - 2017**	0.01	0.46	-0.04	0.24

The extremes of temperature BIAS (WOA18 – SDC) up to 1.2C° are observed in summer months in the upper layer (Figure 26): the negative BIAS in 0-25 m depth range follows by positive BIAS with maximum at about 50 m. The RMSE values are also highest for the same times and depths.

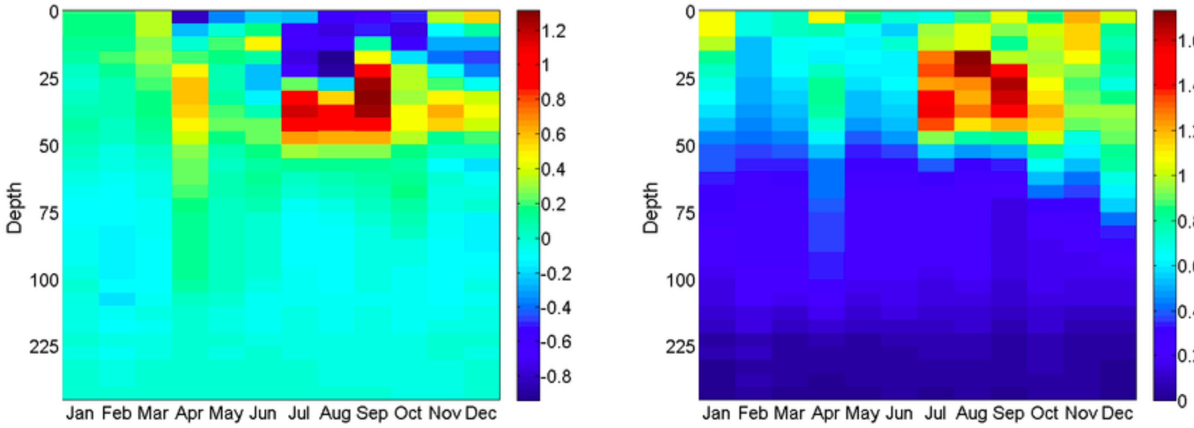


Figure 26 - Hovmoller plots of BIAS (left) and RMSE (right) indexes of WOA18 – SDC monthly temperature fields (upper 300 m).

The maximum of positive salinity BIAS (WOA18 – SDC) up to 0.4 is observed through the year in level 2 = 5m, while the negative values up to -0.2 are observed in depth range 60 – 200m. The RMSE values are highest at the surface (up to 0.9) and in the same depth range 60 – 200m (up to 0.4).

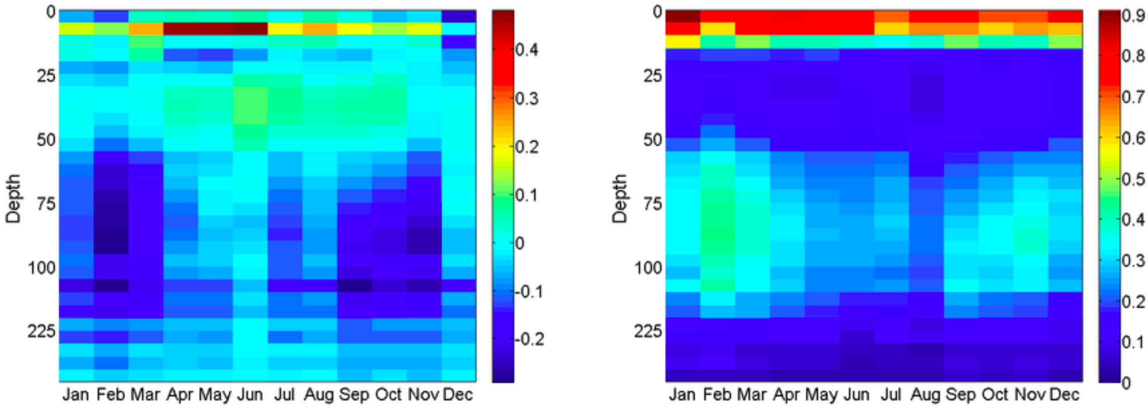


Figure 27 - Hovmoller plots of BIAS (left) and RMSE (right) indexes of WOA18 – SDC monthly salinity fields (upper 300 m).

The BIAS in upper 0-300 layer is confirmed by comparison of average profiles that were produced from climatology for Black Sea internal area for January and July (Figure 28). In January WOA18 and SDC temperature profiles practically coincide, while in July WOA18 temperature is lower in thermocline (0-25m) and higher in CIL (45-55 m). The WOA18 salinity is lower than SDC salinity for the whole halocline layer in both months.

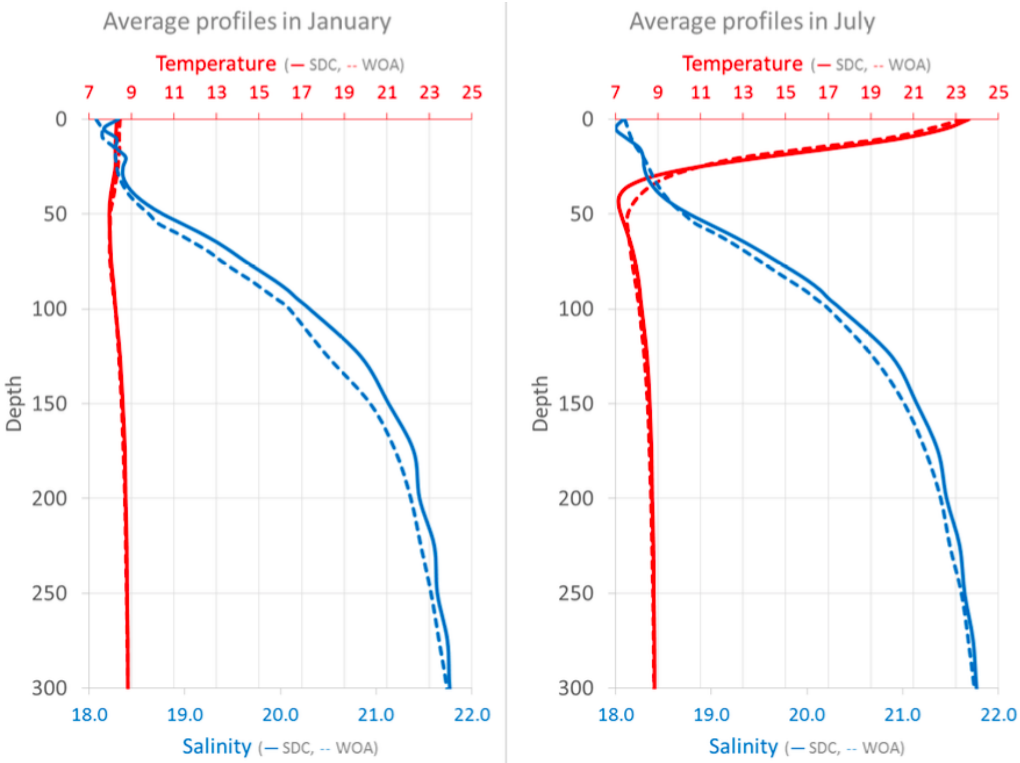


Figure 28 - Average profiles of Temperature and Salinity for area 42 – 44.5N, 31-38E.

6. Technical Specifications

Regional climatologies are delivered as files in NetCDF format. Each file contains four 4d arrays (3 space dimensions + 1 time dimension) named according to the following rule:

- *Parameter_Name* – 4d array for a parameter,
- *Parameter_Name_L1* – … parameter masked using relative error threshold 0.3,
- *Parameter_Name_L2* – … parameter masked using relative error threshold 0.5,
- *Parameter_Name_relerr* – relative error of parameter.

The products are published in Sextant web catalogue and can be accessed from the SeaDataNet portal main webpage https://www.seadatanet.org either by clicking the respective tile (Figure 29) or through the menu.

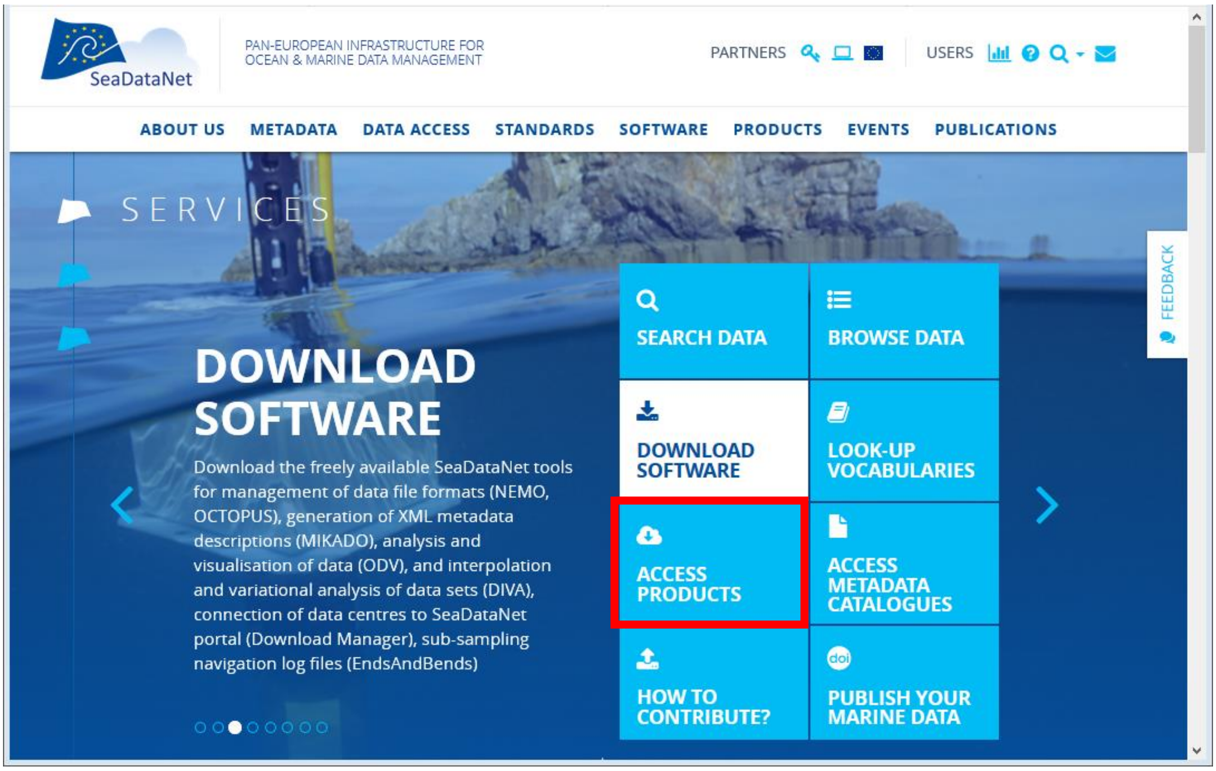


Figure 29 - SeaDataNet portal main page as a gateway to products.

Detailed guidelines on how to access and discover products in Sextant catalogue and how to visualize products is provided in User Manual https://www.seadatanet.org/content/download/3558/file/SDC_UserManual_DataProducts.pdf also accessible from https://www.seadatanet.org/Products/ webpage (Figure 30).

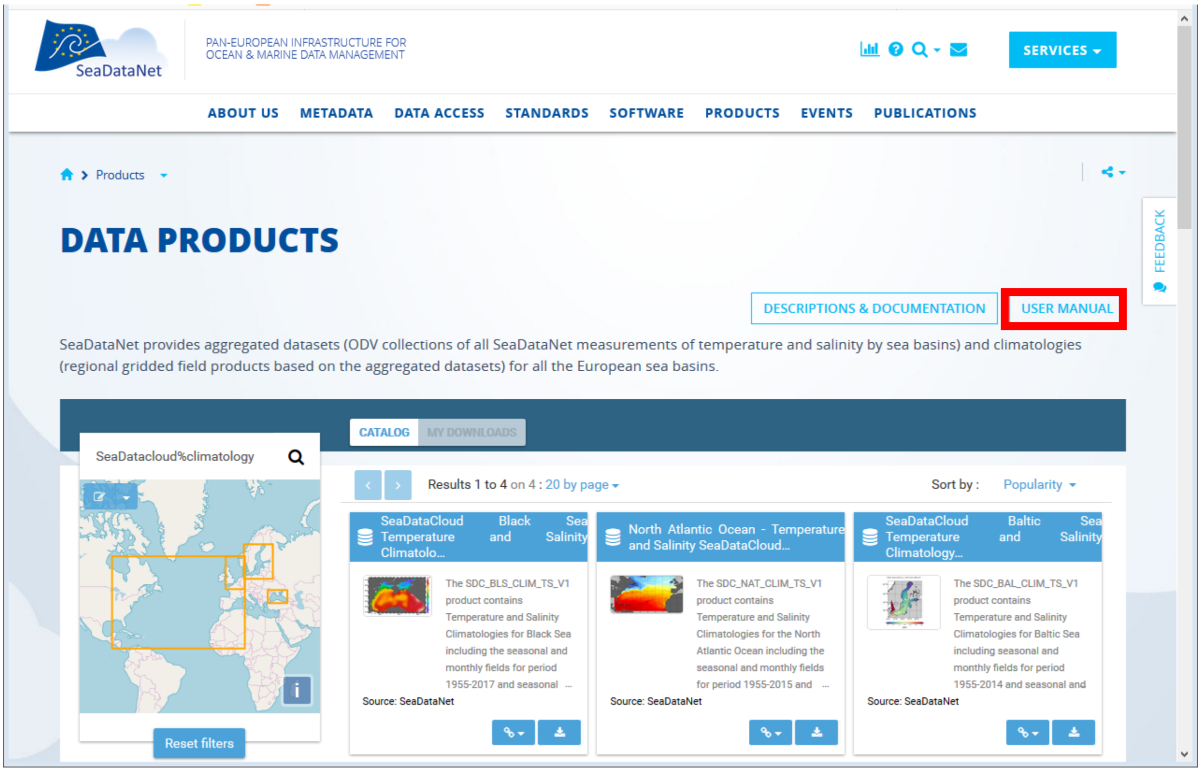


Figure 30 - Data products in Sextant Catalogue at SeaDataNet portals

7. Conclusions

SDC_CLIM_TS_V1 climatologies have been produced with a big effort of all WP11 team. Several gridded fields at different spatial and/or temporal resolution have been generated for the first time, all presents increased vertical resolution. Decadal fields for regional marginal seas have been created for the first time as well.

Two major achievements:

1. Use of new DIVAnd software and team work to debug and improve it.
2. Integration of external sources, analysis on how much they overlap or complement each other and the recognition of data anomalies in the used data sets.
3. PIDocs have been prepared and revised to be annexed to the relative product on the Sextant catalogue.

Both qualitative and quantitative analysis suggested a good consistency of SDC products with WOA and highlighted the added value of regional products compared to the global WOA at lowest spatial resolution. However some minor issues have been identified by regional leaders on the quality of their products and have been reported in the PIDoc section named "Product Usabilty".

The WP11 team already identified possible improvements for the future. Main issues to be addressed in the next SDC climatology releases are:

- Revise each phase of the production chain and optimize it according to the best solutions developed by the WP11 partners;
- Uptake of DIVAnd from all regional leaders and switch to the full 3D version
- Simplify and harmonize the data integration process
- Standardize the duplicate detection process
- Improve the mapping of metadata from WOD and CORA (EDMO code, Instrument/Gear Type, Local CDI ID)
- Better define products' granularity to better fit users' needs and facilitate product accessibility and readiness.
- Implement a procedure to report data anomalies to NOAA and CORIOLIS.

Appendix 1

Naming convention for SeaDataCloud climatological products

Project	Region	Product	Parameter	Version	Product Number	Full Name
SDC	GLO	CLIM	TS	V1	1	SDC_GLO_CLIM_TS_V1_1
SDC	ARC	CLIM	TS	V1	1	SDC_ARC_CLIM_TS_V1_1
SDC	NS	CLIM	TS	V1	1	SDC_NWS_CLIM_TS_V1_1
SDC	BAL	CLIM	TS	V1	1	SDC_BAL_CLIM_TS_V1_1
SDC	NAT	CLIM	TS	V1	1	SDC_NAT_CLIM_TS_V1_1
SDC	MED	CLIM	TS	V1	1	SDC_MED_CLIM_TS_V1_1
SDC	BLS	CLIM	TS	V1	1	SDC_BLS_CLIM_TS_V1_1

Summary of SeaDataCloud climatological products released.

Name	spatial coverage	horizontal resolution	vertical levels	Time coverage	Annual	Seasonal	Monthly	External data sets
SDC_GLO_CLIM_TS_V1_1	Global Ocean	1/4°		1900-2017			**X**	WOD13
SDC_GLO_CLIM_TS_V1_2	Global Ocean	1/4°		2003-2017			**X**	WOD13
SDC_ARC_CLIM_TS_V1_1	45W-70°E 62-83°N	1x1/2°	57 WOD standard levels	1955-1964 1965-1974 1975-1984 1985-1994 1995-2004 2005-2014		**X**	**X**	WOD18
SDC_ARC_CLIM_TS_V1_2	45W-70°E 62- 83°N	1x1/2°	57 WOD standard levels	1955-2014		**X**	**X**	WOD18
SDC_NS_CLIM_TS_V1_1	48.5-62°N 4°W-10°E	1/8°	41 WOD standard levels (up to 700m)	1955-2014	**X**		**X**	WOD2018
SDC_NS_CLIM_TS_V1_2	48.5-62°N 4°W-10°E	1/8°	41 WOD standard levels (up to 700m)	1955-1964 1965-1974 1975-1984 1985-1994 1995-2004 2005-2014		**X**		WOD2018
SDC_BAL_CLIM_TS_V1_1	9-30°E 53-66°N	1/16x1/32°	WOA standard levels (up to 300m)	1955-1964 1965-1974 1975-1984 1985-1994 1995-2004 2005-2014		**X**	**X**	CORA5.2
SDC_BAL_CLIM_TS_V1_2	9-30°E 53-66°N	1/16x1/32°	WOA standard levels (up to 300m)	1955-2014		**X**	**X**	CORA5.2

SDC_NAT_CLIM_TS_V1_1	82W-10°E 10°N-62°N	1/2°	107 WOA standard levels	1955-1964 1965-1974 1975-1984 1985-1994 1995-2004 2005-2015		**X**	**X**	CORA5.1
SDC_NAT_CLIM_TS_V1_2	82W-10°E 10°N-62°N	1/4°	107 WOA standard levels	1955-1964 1965-1974 1975-1984 1985-1994 1995-2004 2005-2015		**X**	**X**	CORA5.1
SDC_NAT_CLIM_TS_V1_3	82W-10°E 10N-62°N	1/2°	107 WOA standard levels	1955-2015		**X**	**X**	CORA5.1
SDC_NAT_CLIM_TS_V1_4	82W-10°E 10N-62°N	1/4°	107 WOA standard levels	1955-2015		**X**	**X**	CORA5.1
SDC_MED_CLIM_TS_V1_1	9.25W-36.5°E 30-46°N	1/8°	WOA18 standard levels	1955-2017		**X**	**X**	CORA5.2
SDC_MED_CLIM_TS_V1_2	9.25W-36.5°E; 30-46°N	1/8°	WOA18 standard levels	1955-1984		**X**	**X**	CORA5.2
SDC_MED_CLIM_TS_V1_3	9.25W-36.5°E; 30-46°N	1/8°	WOA18 standard levels	1985-2017		**X**	**X**	CORA5.2
SDC_MED_CLIM_TS_V1_4	9.25W-36.5°E; 30-46°N	1/8°	WOA18 standard levels	1955-1964 1965-1974 1975-1984 1985-1994 1995-2004 2005-2017		**X**		CORA5.2
SDC_BLS_CLIM_TS_V1_1	27.5-41.875°E; 40.875-47.25°N	1/8°	67 WOA standard levels	1955-1994			**X**	WOD18, CORA5.1
SDC_BLS_CLIM_TS_V1_2	27.5-41.875°E; 40.875-47.25°N	1/8°	67 WOA standard levels	1995-2017			**X**	WOD18, CORA5.1
SDC_BLS_CLIM_TS_V1_3	27.5-41.875°E; 40.875-47.25°N	1/8°	67 WOA standard levels	1955-2017			**X**	WOD18, CORA5.1
SDC_BLS_CLIM_TS_V1_4	27.5-41.875°E; 40.875-47.25°N	1/8°	67 WOA standard levels	1955-1964 1965-1974 1975-1984 1985-1994 1995-2004 2005-2015		**X**		WOD18, CORA5.1

References

*Boyer, T.P., J. I. Antonov, O. K. Baranova, C. Coleman, H. E. Garcia, A. Grodsky, D. R. Johnson, R. A. Locarnini, A. V. Mishonov, T.D. O'Brien, C.R. Paver, J.R. Reagan, D. Seidov, I. V. Smolyar, and M. M. Zweng, 2013: World Ocean Database 2013, **NOAA Atlas NESDIS 72,** S. Levitus, Ed., A. Mishonov, Technical Ed.; Silver Spring, MD, 209 pp.*

Boyer, T.P., O.K. Baranova, C. Coleman, H.E. Garcia, A. Grodsky, R.A. Locarnini, A.V. Mishonov, C.R. Paver, J.R. Reagan, D. Seidov, I.V. Smolyar, K.W. Weathers, M.M. Zweng(2018). World Ocean Database 2018. A. V. Mishonov, Technical Editor, NOAA Atlas NESDIS 87.

Barth, A., Beckers, J.-M., Troupin, C., Alvera-Azcárate, A., and Vandenbulcke, L.: DIVAnd-1.0: n-dimensional variational data analysis for ocean observations, Geosci. Model Dev., 7, 225-241, doi:10.5194/gmd-7-225-2014, 2014.

Data Interpolating Variational Analysis, User Guide, C. Troupin, S. Watelet, M. Ouberdous, D. Sirjacobs, A. Alvera-Azcárate, A. Barth, M.-E. Toussaint & J.-M. Beckers, 2018.

DIVAnd.jl documentation, https://gher-ulg.github.io/DIVAnd.jl/latest/#DIVAnd.jl-documentation-1

The GEBCO Digital Atlas published by the British Oceanographic Data Centre on behalf of IOC and IHO, 2003.

Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M. M. Zweng, C. R. Paver, J. R. Reagan, D. R. Johnson, M. Hamilton, and D. Seidov, 2013. World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 73, 40 pp.

Locarnini, R. A., A. V. Mishonov, O. K. Baranova, T. P. Boyer, M. M. Zweng, H. E. Garcia, J. R. Reagan, D. Seidov, K. Weathers, C. R. Paver, and I. Smolyar, 2018. World Ocean Atlas 2018, Volume 1: Temperature. A. Mishonov Technical Ed.; in preparation.

Zweng, M.M, J.R. Reagan, J.I. Antonov, R.A. Locarnini, A.V. Mishonov, T.P. Boyer, H.E. Garcia, O.K. Baranova, D.R. Johnson, D.Seidov, M.M. Biddle, 2013. World Ocean Atlas 2013, Volume 2: Salinity. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 74, 39 pp.

Zweng, M. M., J. R. Reagan, D. Seidov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, A. V. Mishonov, O. K. Baranova, K. Weathers, C. R. Paver, and I. Smolyar, 2018. World Ocean Atlas 2018, Volume 2: Salinity. A. Mishonov Technical Ed.; in preparation. .

*Szekely Tanguy, Gourrion Jerome, Pouliquen Sylvie, Reverdin Gilles (2016). **CORA, Coriolis Ocean Dataset for Reanalysis**. SEANOE. https://doi.org/10.17882/46219*
Szekely, T., Gourrion, J., Pouliquen, S., and Reverdin, G.: The CORA 5.2 dataset: global in-situ Temperature and Salinity measurements dataset. Data description and validation, Ocean Sci. Discuss., https://doi.org/10.5194/os-2018-144, in review, 2019.

List of Acronyms

Acronym	Definition
ARC	Arctic ocean
BAL	Baltic Sea
BLS	Black Sea
CDI	Common Data Index
CLIM	Climatology
CMEMS	Copernicus Marine Environment Monitoring Service
DATA	Aggregated Dataset
DIVA	Data-Interpolating Variational Analysis (software)
DOI	Digital Object Identifier
EC	European Commission
EDMO	European Directory of Marine Organisations (SeaDataNet catalogue)
GLO	GLobal Ocean
IOC	Intergovernmental Oceanographic Commission
IODE	International Oceanographic Data and Information Exchange (IOC)
MED	Mediterranean Sea
NAT	North Atlantic Ocean
NWS	North West Shelf
ODV	Ocean Data View Software
QC	Quality Checks
QF	Quality Flags
SDC	SeaDataCloud
SDN	SeaDataNet
TS	Temperature and Salinity
WOA	World Ocean Atlas
WP	Work Package