SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

Synthesis on Products and WP10

Antonio Guarnieri
on behalf of S.Simoncelli & M.Tonani (INGV)

OUTLINE

- Refresh of WP10 original duties and time schedule
- new aggregation exercise → V1.1 data collections
- WP10 last year activities
- V1.1 quality assessment results and advance from V1
- Preparation of V1.1 documentation (D10.2 and its user version)
- Restricted data assessment results
- Preliminary work on climatology
- Future Plans→ reviewed duties and time schedule
- Conclusions

WP10 Dow and Deliverables

Common specifications, selected parameters, definition of QC, aggregation and analysis procedures — **Sept 2012**

V1
AGGREGATED DATASET — **Sept 2013**

Examples of data products
gridded fields/ climatological profiles and related std — **Sept 2014**

V2
AGGREGATED DATASET — **Sept 2015**

Delive-rable Number [61]	Deliverable Title	Delivery date [64]
D10.1	Common specifications, selected parameters, aggregation, control and processing procedures	12
D10.2	First release of the aggregated data sets products	24
D10.3	Release of examples of data products	36
D10.4	Final version of aggregated data sets	46

V1 Regional aggregated data sets and D10.2 release have been postponed in order to perform a new aggregation exercise

New aggregation exercise → V1.1 data

StComm and RCs agreed to repeat CDI harvesting and data QC assessment in order to improve the quality of historical data collections

A **new aggregation exercise** was meant to

1. deliver the best aggregated data sets
2. consolidate and refine QC procedures
3. retrieve and analyze **restricted data** for statistical products computation
4. increase the quality of SDN statistical products based on data collections (climatologies)
5. rehinforce and renew the collaboration with MyO: increased quality for SDN, increased quantity for MyO

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

SeaDataNet

WP10 ACTIVITY

Jun2013

Decision to repeat data aggregation

Jul2013

anomalies sent to CDI partners for corrections

Sept2013

deadline for CDI corrections (CDI partners)

QC analysis of V1 AGGREGATED DATASET Presented

Sept 2013

Nov2013-Mar2014

data harvesting (MARIS) and aggregation (AWI)

release of V1.1 to WP10 on Mar10 due to some bugs fixing

V1.1 QC analysis

Restricted data QC analysis

Preliminary study for climatology computation

Internal release of V1.1 AGGREGATED DATASET and delivery of D10.2

Sept 2014

2nd Feedback on V1.1 and restricted data to CDI partners

Oct2014

Internal discussion on dissemination strategy of V1.1 as a real product

WP10 activity focus

1. Debugging of ODV tool (SMHI, Ifremer) in collaboration with AWI during the data aggregation phase
2. Analysis V1.1 regional data collections (1900-2012) and comparison with V1
3. Analysis of restricted data sets
4. Preliminary work on climatology: DIVA settings, time resolution (seasonal/monthly), background field definition
5. Preparation of V1.1 data documentation (D10.2, User Guide)

SeaDataNet

Debugging activity (Jan-Mar 2014)

First version of the V1.1 aggregated dataset has been provided by AWI at the end of Jan2014. Working on this new version, bugs (like mix of several profiles in a same profile, gap in metadata rows) have been detected and have been fixed by a new version of ODV and a new version of aggregated dataset delivered in Mar2014.

An additional improvement was the new parameter Pressure in addition to Depth to identify levels.

Jan 24, 2014 → **1st version aggregated dataset was provided by AWI**

Complex set of iteration and bug fixing processes including software bugs/fixings, new releasea, metadata problems, labelling problems in metadata, etc…

Mar 17, 2014 → **new regional data collections provided to SeaDataNet RCs together with restricted data sets to be used only for SeaDataNet products creation**

List of iterations/actions undeertaken for the V1,1 Debugging

- 24 Jan: **AWI** provided the 1st version aggregated dataset (ODV 4.5.7) – split into regions **(Ifremer)**
- 29 Jan: new release ODV 4.6.0
- 4 Feb: Feedback on ODV 4.6.0 about problems with log file (no way to record changes on QC) -> use ODV 4.5.7 to edit log **(Ifremer)**
- 6 Feb: Feedback about ODV bug (Temp and TEMP not unique labels) fixed with ODV 4.6.0 but the 2nd dataset was created with ODV 4.5.7 **(SMHI)**
- 13 Feb: new version ODV 4.6.1 to test (Export→Stations History to get changes on QC)
- 18 Feb: official version ODV 4.6.1
- 20 Feb: ODV 4.6.1 problem with Metadata **(SMHI)**
- 25 Feb: problem with salinity → 3 data files being merged into 1 profile **(SMHI)**
- 27 Feb: gaps in temperature profile levels **(Ifremer)**
- Problems between versions due to creation in ODVCF5 or ODVCF6 format collections (ODVCF5 = no history feature, all edit logs are written to the collection log file)
- 3 Mar: AWI proposed a new aggregated dataset and a new release of ODV 4.6.2
- 10 Mar: new aggregated dataset → collections are ODVCF5 and can be opened by ODV 4.5.7 as well as ODV 4.6.0 or later
- **17 Mar: new regional data collections provided to SeaDataNet RCs together with restricted data sets to be used only for SeaDataNet products creation**

Analysis of V1.1 data collections: comparison with V1

V1.1 data collections show a general data population increase due to the insertion of new data especially in the Atlantic region

	V1	V1.1	increase %
Atlantic (#data)	431974	1049547	143
Baltic Sea (#points)	8900000	11700000	31
Black Sea (1990-2013 #stations)	21068	23142	10
Mediterranean Sea	136828	169438	24
Arctic	407711	445281	9.
North Sea	n.a.	n.a.	n.a.

Results of Analysis of V1.1

Atlantic: due to the large increase of data, **anomalies on QF are still observed**. Data with **QF zero are still in the dataset. Some files only contain depth and pressure.** A list of those anomalies has to be sent to the CDI Partners to allow them to make corrections on their QF.

Baltic Sea: The **quality** of the data set from the Baltic Sea **is high**. Less than **1%** of all data were **flagged as suspicious** after a thorough quality control. Data are **well distributed both spatially and temporally**. This makes a good basis for monthly climatologies.

Black Sea: data set contains **large number of not QC data and erroneous data flagged as good.** The dataset contains a lot of **duplicates** too. In order to be utilized in oceanographic and ecological applications the dataset should undergo thorough quality assessment.

Mediterranean Sea: V1.1 data quality is better than V1 thanks to the QC strategy implemented, but there are **still bad data flagged as good, data are on land and with wrong longitude format** (360° instead of ±180°)

Arctic Waters: V1.1 data content increased due to 30000 new CTD casts (9%). There are **still bad data flagged as good** and **some unchecked data (QF0) appear as "good" (CF1).**

V1.1 documentation

- A template for the regional QC reports to be included in D10.2 (*) was sent by WP10 leader;
- D10.2 has to be homogenized before official release (these days);
- User Guide for V1.1 products will be finalized from D10.2;
- The required information for V1.1 insertion in SDN catalogue has to be prepared.

(*) D10.2: *First Release of the Aggregated DataSet Products*

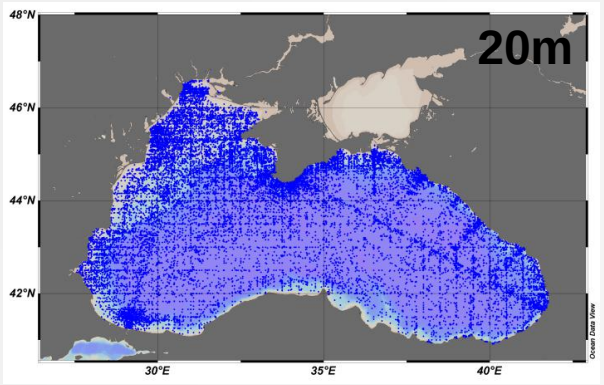
Analysis of restricted data

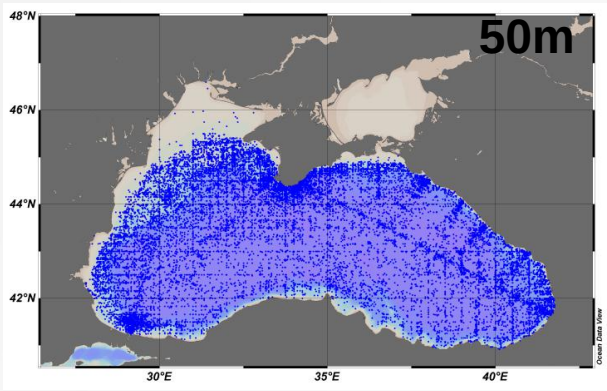
	V1.1	Restricted	%
Atlantic (#data)	1049547	28318	3
Baltic Sea (#points)	11700000	65000	1
Black Sea (1990-2013 #stations)	23142	4287	19
Mediterranean Sea	169438	28690	17
Arctic	445281	245	~0%
North Sea	n.a.	n.a.	n.a.

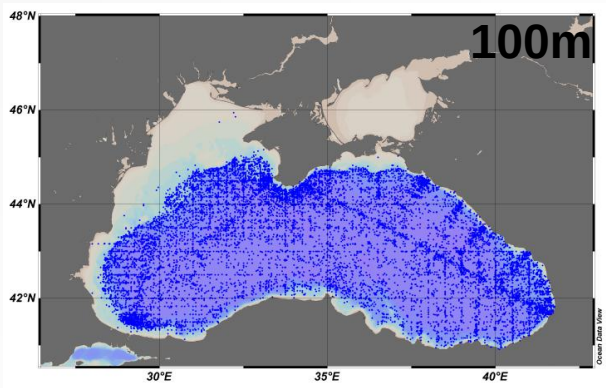
Restricted data increased V1.1 data collections with different percentages. In particular Black Sea and Mediterranean regions present still a cospicuos percentage of restricted data, that are fundamental for climatology computation. Mediterranean restricted data that are not included in the data harvesting process are still a lot (estimate of another 20% from web portal).

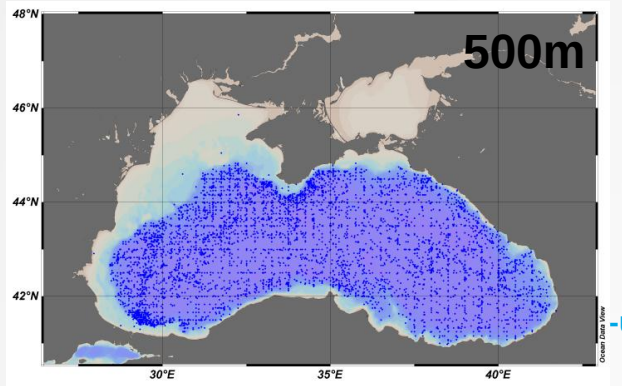
Preliminary work on climatology

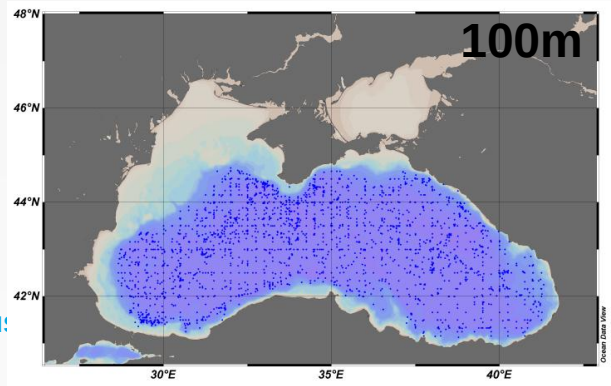
The **Black Sea** dataset can be used for computing climatologies: **data availability drastically decreases with depth** as shown on density of data coverage for selected IODE depth levels. It is planned to calculate **seasonal climatic maps of T and S** for depth levels 0, 10, 20, 30, 50, 75, 100, 150, 200, 250 m and **annual maps** for depth levels 300,400, 500, 600, 800, 1000, 1200, 1500, 2000 m.

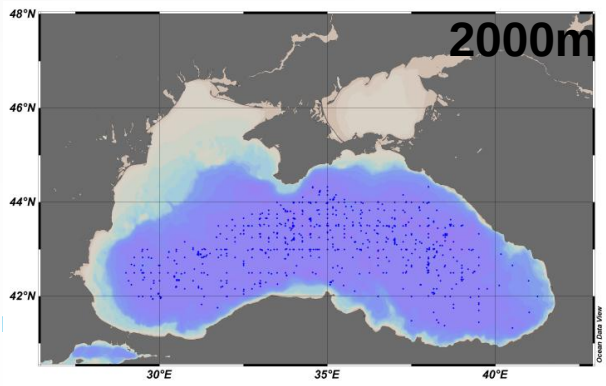








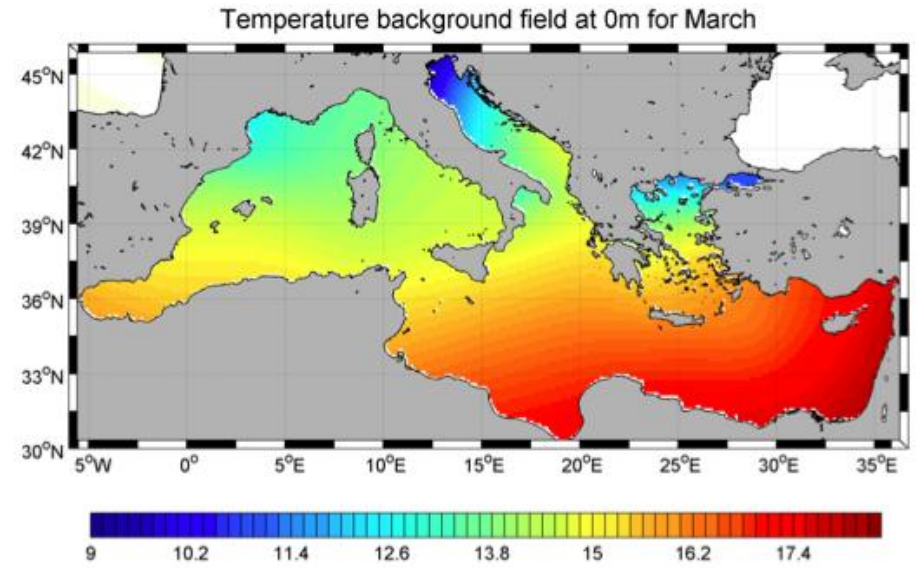




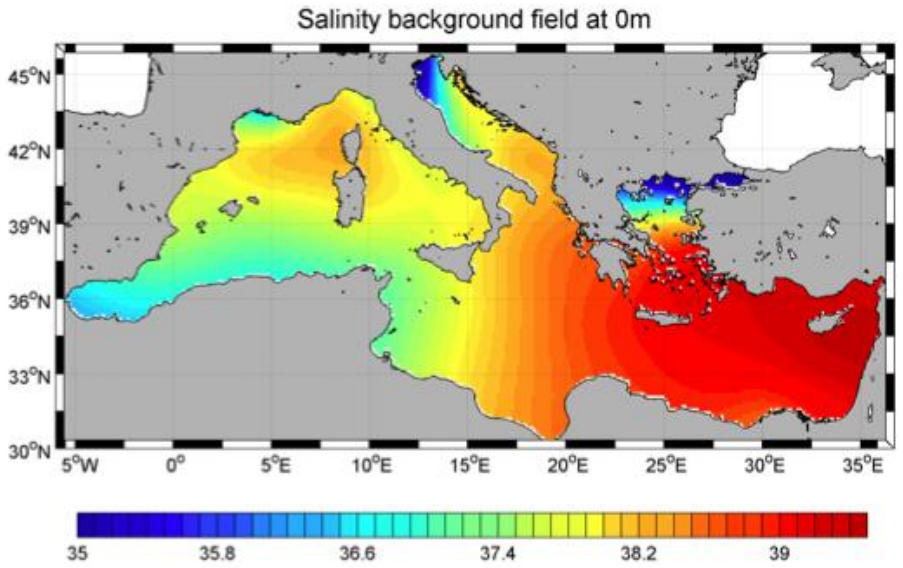
Preliminary work on climatology

Mediterranean Sea: climatologies at 1/8° horizontal resolution were produced for 33 IODE standard levels.

Background fields: - T monthly seminorm splines reconstructions
- S annual averages for salinity



(a) Temperature backgound at 0m



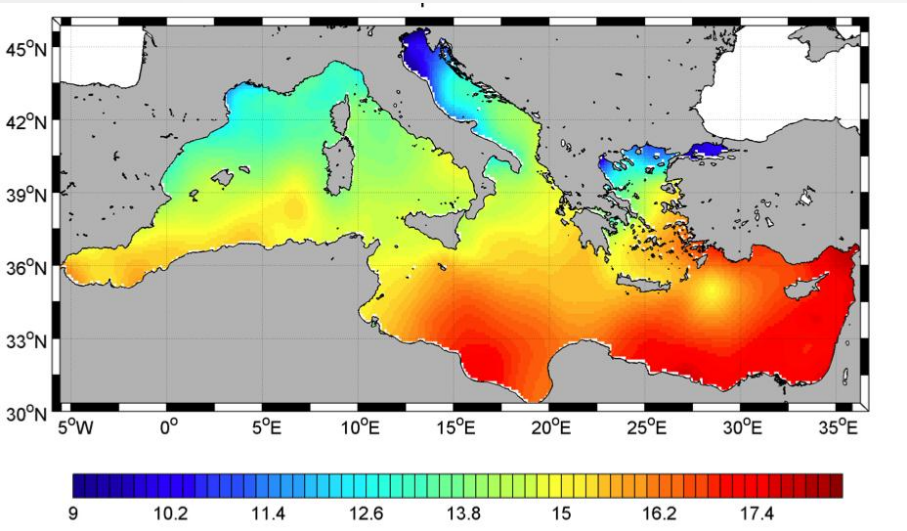
(b) Salinity backgound at 0m

Preliminary work on climatology

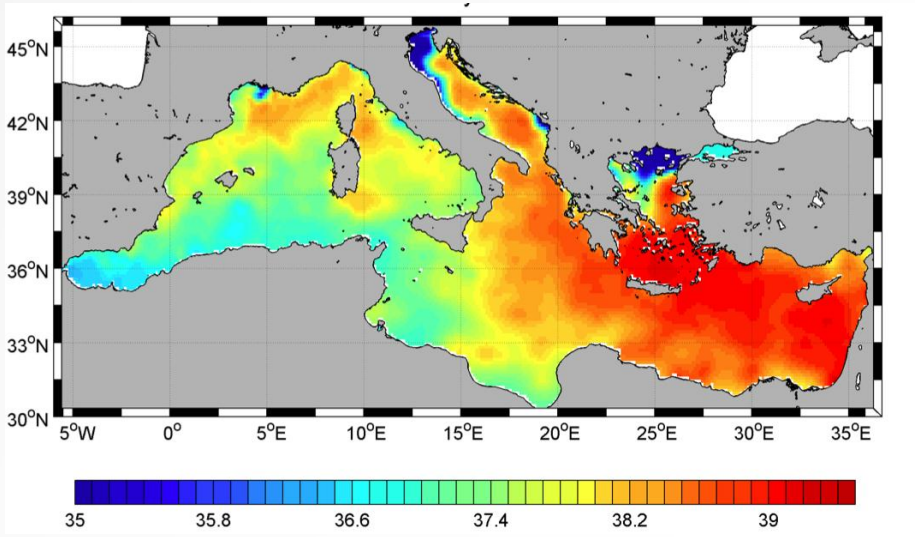
Mediterranean Sea: climatologies at 1/8° horizontal resolution for 33 IODE standard levels

Background fields: - T monthly seminorm splines reconstructions
- S annual averages for salinity

Consistency and quality of climatological fields have been checked versus previous SDN climatology



March Temperature Climatology at the surface

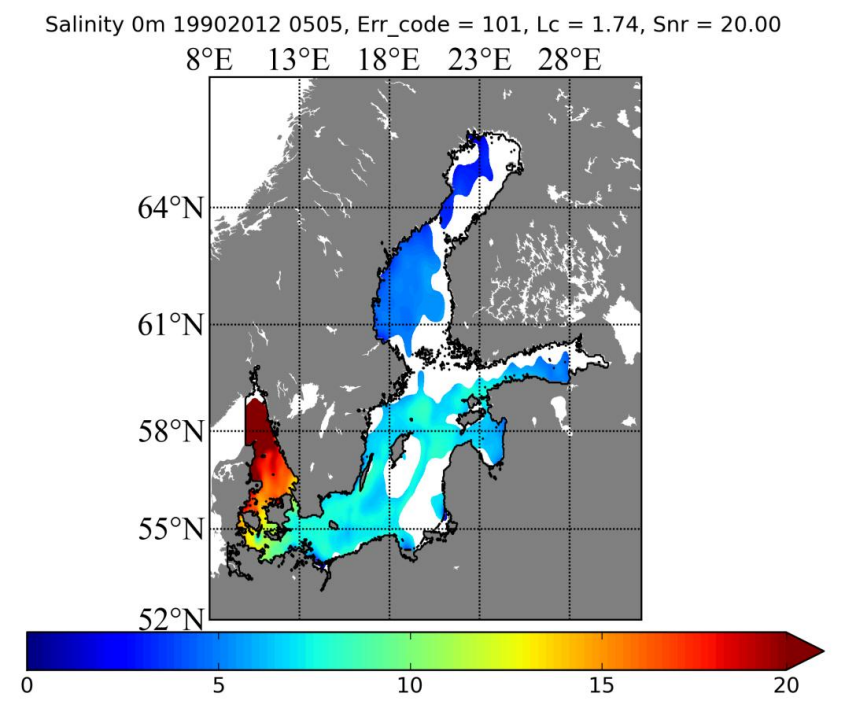


Seasonal Salinity Climatology at the surface (Apr, May,Jun)

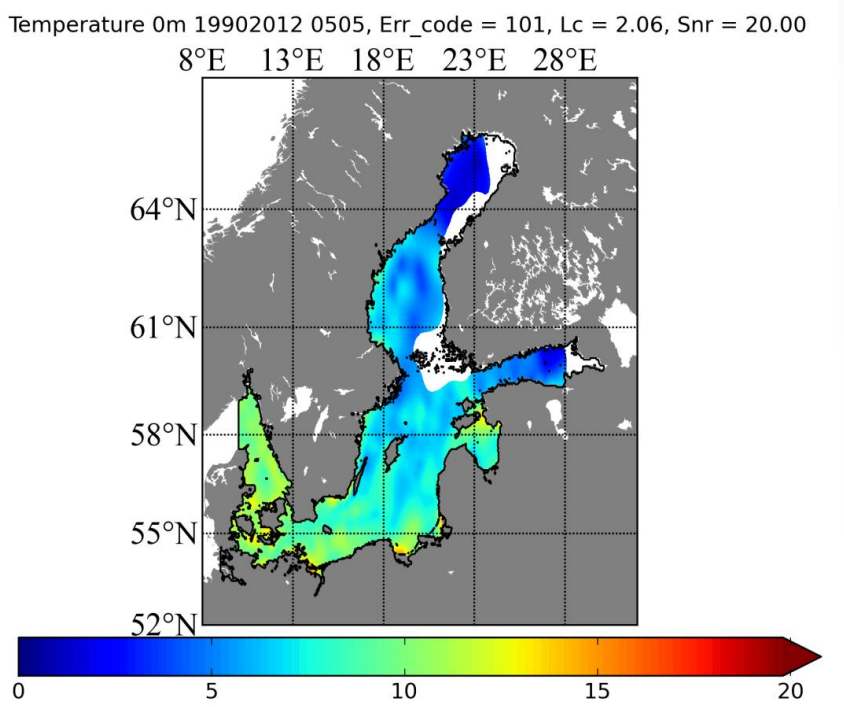
Preliminary work on climatologies

Baltic Sea: Some DIVA **tests** have been done but still need to tune DIVA parameters and c**reate seasonal background fields**

Salinity



Temperature



Salinity and Temperature at the surface in May with masked errors

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

Future Plans

Release of D10.2 and internal distribution of V1.1 aggregated dataset — **Sept 2014**

1. Internal discussion to be opened and finalized on dissemination strategy of V1.1
2. Joint discussion with MyO community to decide on the release strategy of the joint product (mid Oct-mid Nov)

2nd Feedback on V1.1 and restricted data to CDI partners — **Oct 2014**

Dec 2014 CDI partners deadline for data check and correction
Jan2015(*) new harvesting procedure (MARIS)
Feb2015 (*) new aggregation procedure (AWI)

1. Internal **release of data products** (climatologies)
2. **(*)** Release of aggregated to WP10 — **Mar 2015**

QC analysis on V2 data collections validation of climatologies

V2 release of AGGREGATED DATASET and data products in SDN catalogue — **Sept 2015**

(*) subject to the outcomes of the joint meeting

SeaDataNet

QC Procedure for V2

Central CDI

data harvesting and aggregation

new data entries or corrections

QC analysis at regional level by **SDN RCs** using ODV

reply to data anomalies

Dec2014 CDI partners make corrections in central CDI and report on corrections

RCs provide list of V1.1 and restricted data anomalies

delivery of anomalies files

Oct2014

analysis of anomalies and organization by EDMO code

QC procedure has been designed to be iterative in order to facilitate the update and improvement of SDN database content→ V2

Conclusions

- All V1.1 data collections contain more data than V1 version
- Quality assessment highlighted that there are still anomalous data and data not checked (QF=0), but dataet is ady from quality point of view;
- Restricted data present an overall good quality but they represent for some regions a too large percentage (Med and Black Seas)
- **instructions on corrections of metadata and data and elimination of duplicates should be sent to data providers for applying as soon as possible otherwise V2 will contain the same errors**
- **RCs agree on the need of a second round of feedback to data providers for V2**
- Data collections should be released in their original format without modifications by RCs, at least at this stage. After Climatology release and data citation through DOI, V2 might have all the RCs corrections
- RCs partecipated to WP activities (except for North Sea)
- QC reports have been sent to the coordinator and has to be homogenized
- User guide has to be adapted starting from D10.2 information
- Dissemination through SDN web catalogue requires some specific description → delivery to external users depends on the amount of work required and on the decisions taken in the internal discussion (Oct 2014)