

SeaDataCloud Temperature and Salinity data collections

S. Simoncelli (INGV)

simona.simoncelli@ingv.it

C. Coatanoan (Ifremer), V. Myroshnychenko (METU), Ö. Bäck (SMHI), H. Sagen (IMR), S. Scory (RBINS), N. Pinardi (UniBO), R. Schlitzer (AWI), A. Barth (ULG)



Marine Data Value Chain

DISCOVER

COLLECT

ORGANIZE

> INTEGRATE

ANALYZE

VALIDATE & REPORT

VISUALIZE & DISSEMINATE

USE

SUPPORT ACTIVITIES

Coordination and Communication

System Monitoring

Technology Infrastructure

Service Development

Create Knowledge,
Derive Products,
Develop
Applications and
Services,
Drive Predictive
Models,
Rapid
Environmental
Assessment

INNOVATION

EFFECTIVE &
TRANSPARENT
DECISION
SUPPORT

BLUE SUSTAINABLE GROWTH

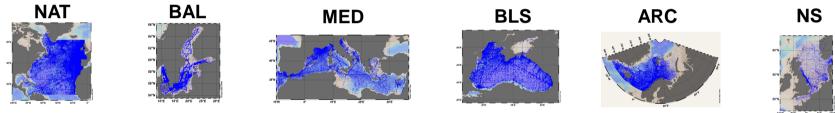
FAIR guiding data management principle and linking data approach
Science support and supervision
Up to date QC procedures
User driven and provider/producer appreciative



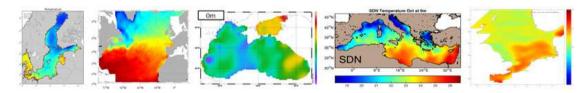
SeaDataCloud Products

GOAL: to provide the best data products from SeaDataNet at regional and global scale and serve diverse user communities (op. oceanography, climate, marine environment, institutional, academia)

 Aggregated data sets EU marginal seas → historical temperature and salinity data harvested from the central CDI and validated by regional leaders



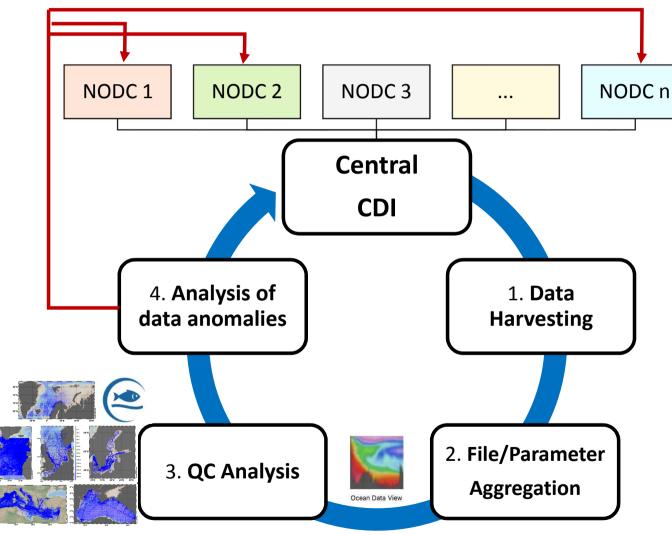
2. Climatologies → gridded fields obtained through DIVA mapping tool and representing the climate of the ocean at regional and global scale



3. New data products → multi-platform and multi-disciplinary approach combining in situ (e.g. gliders, Argo, ships, drifters, fixed platforms) and remote sensed observations, Ocean Monitoring Indicators for tracking ocean mechanisms and/or climate modes and trends



Quality Control Strategy (QCS)



GOAL: to improve the quality of SeaDataNet database content and create the best data products

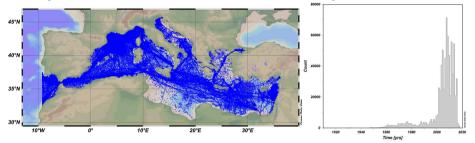
→ iterative approach to facilitate the upgrade of the database and versioning of data products



Quality Control Analysis

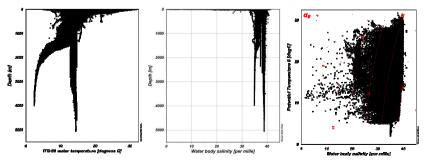
Common guidelines

- 1. Spatial data distribution
- 2. Temporal data distribution (annual, seasonal and monthly)
- 3. Quality Flags statistics
- 4. Scatter plots of good/probably good (QF1/QF2) observations
- 5. Gross range check
- 6. Analysis of QFO data (not checked) to disclose good data
- 7. Visual check to identify wrong profiles (spikes, outliers)
- 8. Identification of stations on land
- 9. Identification of wrong/missing data (time, measurements)
- 10.Stability check



par	# stations	%	# samples
total	739784		
Т	737102	99,6	41223938
S	667232	90,2	28518744
TS	665388	89,9	28119926

%	QF0	QF1-2	Q3-4
dpt	3.0→0	96.9→99.8	0.1
Т	2.7 → 0	97.0→99.8	0.3
S	4.5→0	94.6→99.2	0.9
dpt&T&S	3.0	94,4	0.3





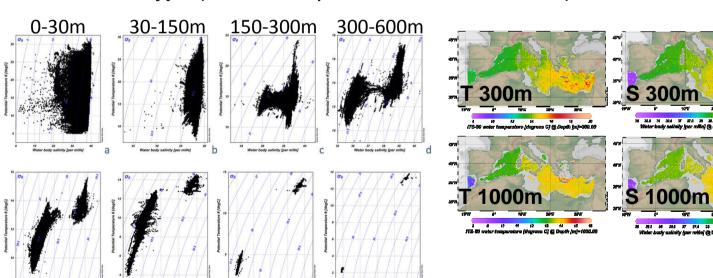
Quality Control Analysis

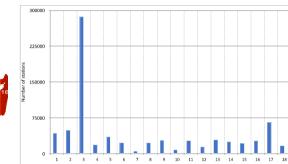
Specific checks per

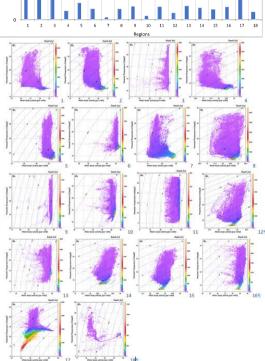
- areas (similar hydrodynamic characteristics)
- layers (surface, intermediate, bottom)
- time periods (decades or specific periods)

600-1000m 1000-2000m 2000-3000m 3000-5500m

Instrument type (consistency issue of historical data)









Metadata Analysis

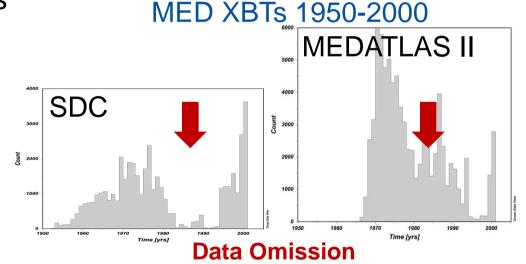
Ocean Data View

New data distributors/originators statistics → QC filtering by EDMO code → detection of systematic (format, flagging) errors

New instrument type statistics -> analysis of monitoring space-time

coverage → detection of data omissions

Instrument/Gear Type	# stations	%
CTD	52031	7
bathythermograph	56558	8
discrete water sampler	32258	4
thermosalinograph	555269	75
thermistor chains	22	0
continuous water sampler	1577	0
salinity sensor; water temperature sensor	19852	3
salinometers	100	0
salinity sensor	143	0
water temperature sensor	1	0
none info	21973	3



!!!Actions needed to ingest missing XBTs!!!

!!!Actions needed to complete crucial metadata information!!!



Metadata Analysis

Ocean Data View

- New data distributors/originators statistics → QC filtering by EDMO code → detection of systematic (format, flagging) errors
- New instrument type statistics → analysis of monitoring space-time coverage → detection of data omissions

Instrument Info	Probe	# stations	%
"		26162	46
'SDN:P01::ADEPZZ01 SDN:L22::TOOL0262 SDN:L33::011 SDN:P01::TEMPET01 SDN:L22::TOOL0262'	T-5 XBT	1239	2
'SDN:P01::ADEPZZ01 SDN:L22::TOOL0263 SDN:L33::041 SDN:P01::TEMPET01 SDN:L22::TOOL0263'	T-7 XBT	9995	18
'SDN:P01::ADEPZZ01 SDN:L22::TOOL0263 SDN:L33::041 SDN:P01::TEMPPR01 SDN:L22::TOOL0263'	T-7 XBT	14	0
'SDN:P01::ADEPZZ01 SDN:L22::TOOL0435 SDN:L33::001 SDN:P01::TEMPET01 SDN:L22::TOOL0435'	T-4 XBT	16732	30
'SDN:P01::ADEPZZ01 SDN:L22::TOOL0592 SDN:L33::710 SDN:P01::TEMPET01 SDN:L22::TOOL0592'	XCTD-2	6	0
'SDN:P01::ADEPZZ01 SDN:L22::TOOL0718 SDN:L33::061 SDN:P01::TEMPET01 SDN:L22::TOOL0718'	T-10 XBT	2126	4

!!!Actions needed to complete crucial metadata information!!!



Data and Metadata Omissions: MED XBTs

	EDMO	# stations	%
OGS	120	491	1
ENEA	136	10488	19
HCMR	269	869	2
Ifremer	486	13549	24
SHOM	540	30876	55
МНІ	727	1	0

ENEA, OGS and Ifremer have been informed:

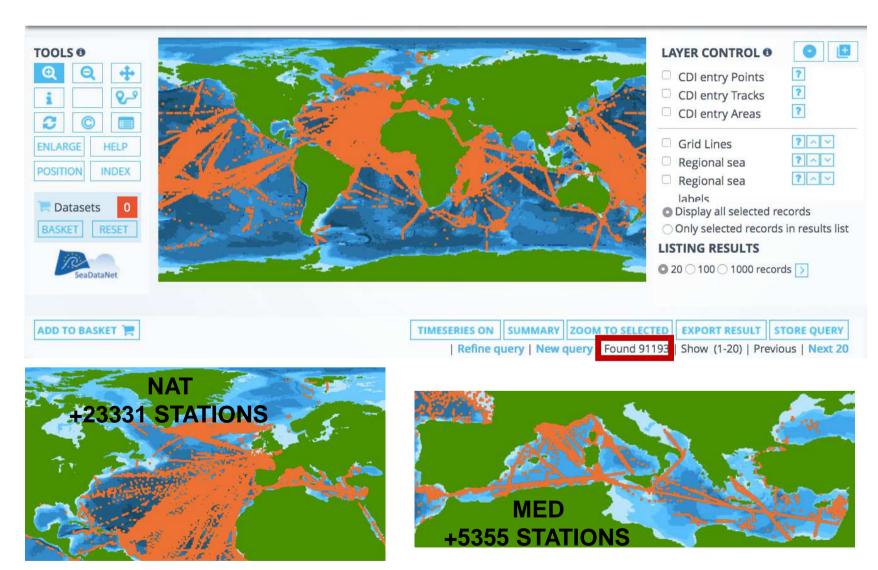
- ENEA will insert all the probe info available and missing metadata
- Ifremer checked and added additional available metadata and ~91k French XBTs
- OGS will insert missing probe and instrument type info

Filling data and metadata omissions highly improves the quality of the infrastructure content and increases users' confidence Reprocessing at the data center level allows

- → to secure and preserve crucial historical info (international approach of preserving the original data)
- →to apply the latest XBT bias correction (ODV development)



New French XBTs





Results: database progresses

Cruises		Stations		Samples (data)				
SDN2_V2	SDC_V1	±%	SDN2_V2	SDC_V1	±%	SDN2_V2	SDC_V1	±%
			Med	iterranear	sea			
			212887	734957	+245,%	26625173	42294299	+59%
				Black Sea				
1723	2284	+32.6%	96487	137723	+43%	2696215	4240346	+57%
				Arctic Sea				
1075	1956	+82%	266291	731286	+175%	19681474	24203161	+23%
				Baltic Sea				
						11100238	13780801	+24%
North Atlantic								
			1807266	9091773	+403%			
	North Sea							
			115 596	162 452	+41%	6670529	7817193	+17%



Product Information Documents (PIDoc)

PIDocs contain all specifications and descriptions of:

- Product's characteristics (format, space-time coverage, resolution)
- Quality (validation methodology and results)
- Product's usability
- data distributors and data originators list (add statistics)
- instrument type statistics

PIDocs have DOI and are available through the product landing page

- → Big effort to produce/revise and publish PIDocs
- → Major improvement to increase user confidence and products uptake



Results

- Data population statistics per sea basin show a progressive increase of available data
- Data quality also improved thanks to the introduction of additional checks by regional experts (sub-regions, depth layers, iso-surfaces)
- QF statistics after QC present very high percentages of good data (QF1,2) ~99% MED; 98-99% BLS; ~99% ARC; ~99% BAL; 98-99% NS 96(S)-99% NAT
- metadata statistics about data distributors/originators highlights systematic (format, flagging) errors and allows the monitoring the EU data sharing landscape but also fair acknowledgment to providers
- instrument type statistics highlighs omissions and suggests the need of systematic check and data reprocessing at the data centers level



SDC_DATA_TS_V1 release (June 2018)



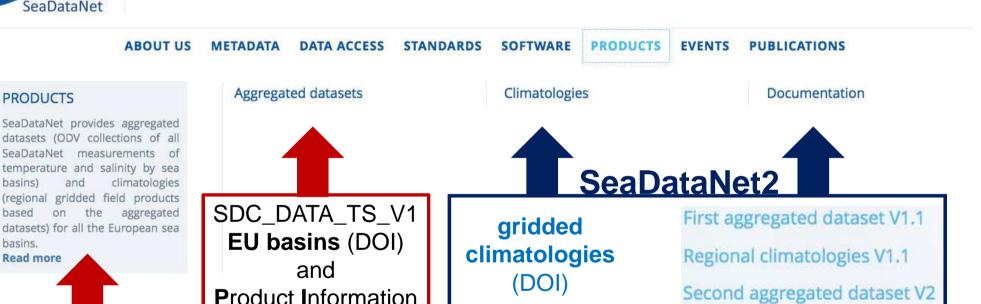
catalogue

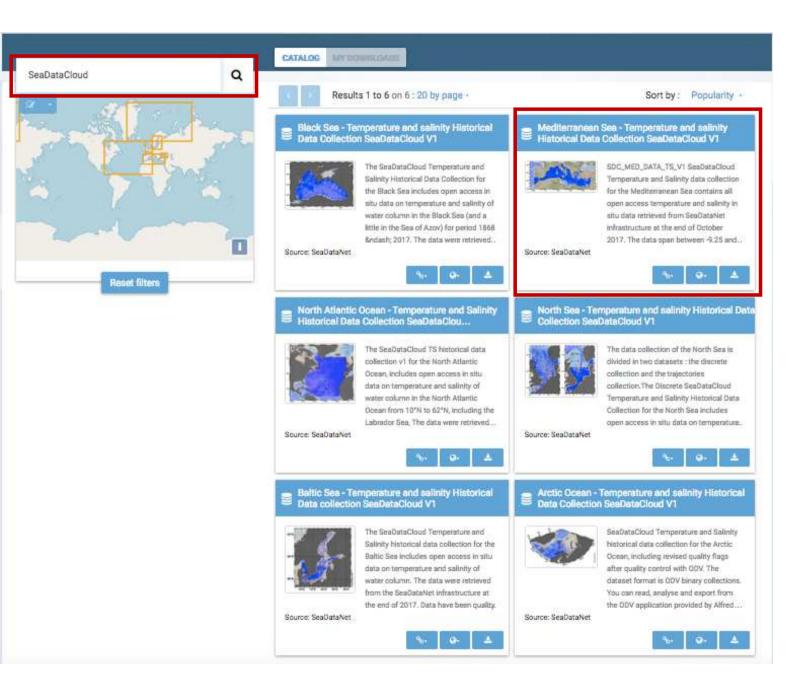
PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

Documents

(DOI)







Products Catalogue



Products Catalogue

Ifremer.

TXT. RIS. XLS

IFREMER / IDM / SISMER

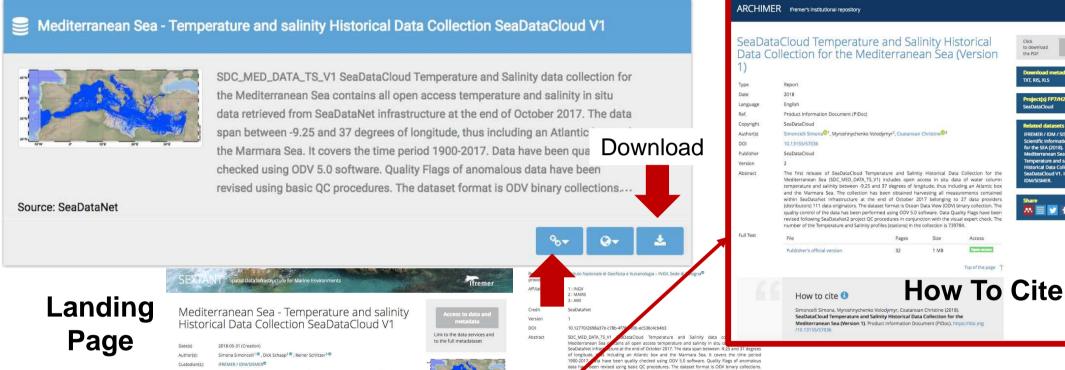
Mediterranean Sea -Temperature and salinity

Scientific Information System for the SEA (2018).

Historical Data Collection SeaDataCloud V1: IFREMER

M ≡ **y** f **H**

Access



Marine Laboratory³⁷, Proudman Oceanographic Laboratory³⁷, CNR, Istituto di Scienze Marine (Sezione di Venezia - ex IBM)³⁸, OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale), Division of Oceanography³⁸, SACLANT Undersea Research Centre (SACLANTCEN)³⁸, CNR, Institute of Marine Science (ISMAR) (Triested CNR Istituto per lo Studio della Dinamica della Grandi Masses Marine Science (SMAR) (Trieses)², C.NR, Istutus per lo Studio della Diamica delle Grand Masse², Institute of Marine Science S.S. of Let (1976). ENTA CENTRO L'ASSERTION CONTROLLA DIAMICA SCIENCE ADMINISTRATION CONTROLLA DIAMICA SCIENCE ADMINISTRATION CONTROLLA DIAMICA SCIENCE ADMINISTRATION CONTROLLA DIAMICA SCIENCE (SMAR) - Altoposity of CRNS. Listyl Letter of Marine Science (SMAR) - Altoposity CRNS. Listyl Letter of Marine Science (SMAR) - Altoposity CRNS. Listyl Letter of Almospheric Science and Climate (SMAP) - Biologia², C.NR, Listyllace of Almospheric Science (SMAR) - Altoposity CRNS. Controllar Controllar Controllar CRNS. Controllar Control Anton Domin Or Squies' Warmer Bology Loboustory of reserving recommendation of Conference of Confere VILLEFANCHIE (LOV) / OO/®, MUSEUM NATIONALO HISTORIC NATURELLE / LABORATORIE
DO DOCAMOGRAPHE PHYSIQUE®, COM, Physical and Biogeochemical Georgapity Laboratory
(LUMINNY). MUNESTI O'REPRIONANY CEEPE. Developmental PRIEMBER / RBB Depayment / Biogeochimical
microproductional Reach Unit (Nature). Providence / RBB Depayment / Biogeochimical
microproductional Reach Unit (Nature). Developmental biogeochimical biogeo Oceanography[®], P.P.Shirshov Institute of Oceanology, RAS[®], National Institute of Fisheries Research (INRH)[®], Institute of Marine Sciences, Middle East Technical University[®], Institute of Oceanograp and Fisheries Res. Center for marine research - Rudjer Boskovic Institute[®], International Ocean Institute - Malta Operational Centre (University Of Malta) / Physical Oceanography Unit® Malta

Originator(s):

Is cited by Myroshnychenko Volodymyr Coatanoan Christine (2018). SeaDataCloud Temperature and Salinity Historical Data Collection

for the Mediterranean Sea (Version 1), Product Information

column. Water body salinity. ITS-90 water temperature, Mediterranean Sea, Eastern Basin, Adriatic Sea, Ionian Sea, Mediterranean Region, Mediterranean Sea, Alboran Sea, Strait of Sicily, Ligurian Sea, henian Sea, Sea of Marmara, Aegean Sea, Mediterranean Sea, Western Basir The data used as input for this product have been extracted from the SeadataNet Download Service: http://www.seadatanet.org/Data-Access/Common-Data-Index-CDI SeaDataNet Usage Agreement: If you use SeaDataNet data in your publications please include the following in the Acknowledgements: Data were provided through SeaDataNet Pan-European

ead, analyse and export from the ODV application provided by Alfred Wegener institute a

nographic geographical features, Temperature of the water column, Salinity of the water

infrastructure for ocean and marine data management (https://www.seadatanet.org)

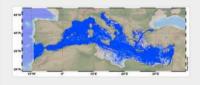
How to cite • How To Cite



Products Catalogue



Mediterranean Sea - Temperature and salinity Historical Data Collection SeaDataCloud V1



SDC_MED_DATA_TS_V1 SeaDataCloud Temperature and Salinity data collection for the Mediterranean Sea contains all open access temperature and salinity in situ data retrieved from SeaDataNet infrastructure at the end of October 2017. The data span between -9.25 and 37 degrees of longitude, thus including an Atlantic box and the Marmara Sea. It covers the time period 1900-2017. Data have been quality checked using ODV 5.0 software. Quality Flags of anomalous data have been revised using basic QC procedures. The dataset format is ODV binary collections....

Source: SeaDataNet

- SDC_MED_DATA_TS_V1
- Water_body_salinity
- + Add all 3 layers to the map







Restricted data analysis

Restricted/Unrestricted ratio

- PIDocs are available
- Metadata stats →
 to monitor continuously
 the amount of restricted
 data, the providers

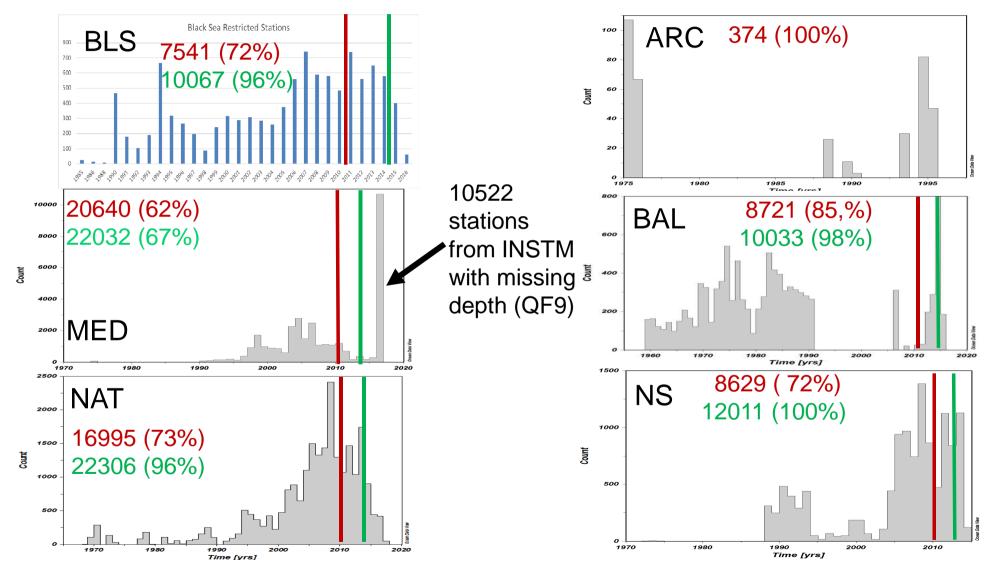


GOAL unlock old data and avoid stagnation

	STATIONS			
	unrestricted	restricted	%	
NAT	9091773	23217	0,3	
BAL		10221	1,0	
MED	734957	33022	4,5	
BLS	137723	10528	7,6	
ARC	731286	374	0,1	
NS	1385289	12011	0,9	



Unlock restricted data before 2010? 2014?





Major providers of restricted data

Region	EDMO	Originator	% st
MED	1232	Institut National des Sciences et Technologies de la Mer (INSTM)	51
BLS	723	State Oceanographic Institute (SOI)	56
ARC			
BAL	193	Institute of Meteorology and Water Management National Research Institute, Maritime Branch in Gdynia (IMWM MB)	81
NAT		IEO	66
NS	2135	Marine Scotland Science	37

Coming soon \rightarrow monitoring of restricted/unrestricted rationper data provider to assure that the balance is reasonable



TRAINING and DISSEMINATION

DIVA training 3-6 April 2018 \rightarrow wide participation, all RC were present Very good and efficient course

SDC_WP11_D11.14_StrategyForTrainingActivities.pdf

Modified on: 22 December, 2017 By: Michele FICHAUT

SDC_WP11_D11.15_OutcomeOfTrainingActivities.pdf

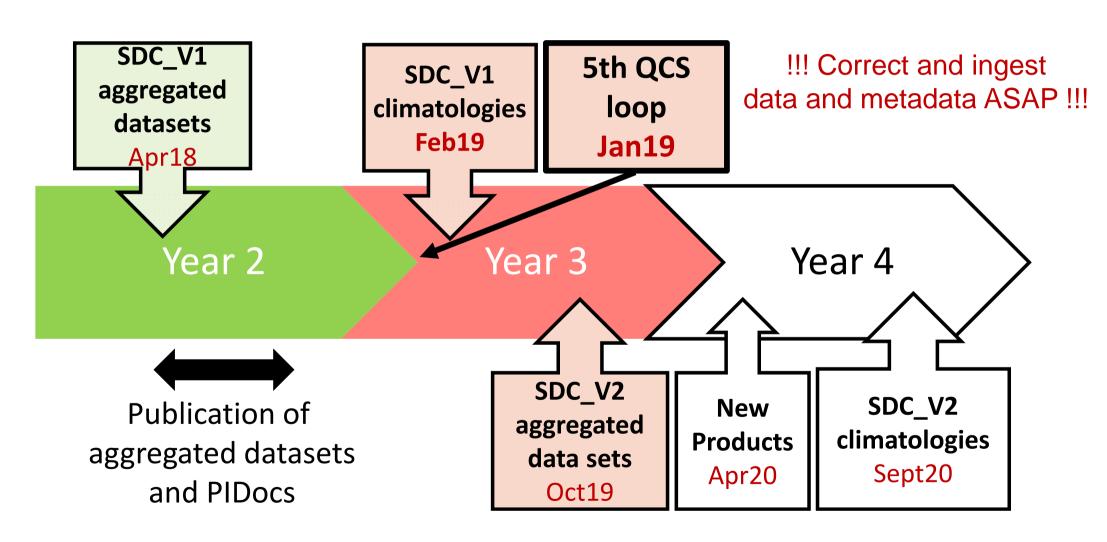
Modified on: 20 April, 2018 By: Michele FICHAUT

→Importance Quality Control SDC WP11 Introduction Prestentation to the 1st SDC training workshop (Serge, Simona, Christine)

WP11 promoted and presented SDC in many workshop and conferences (WP4 presentation)



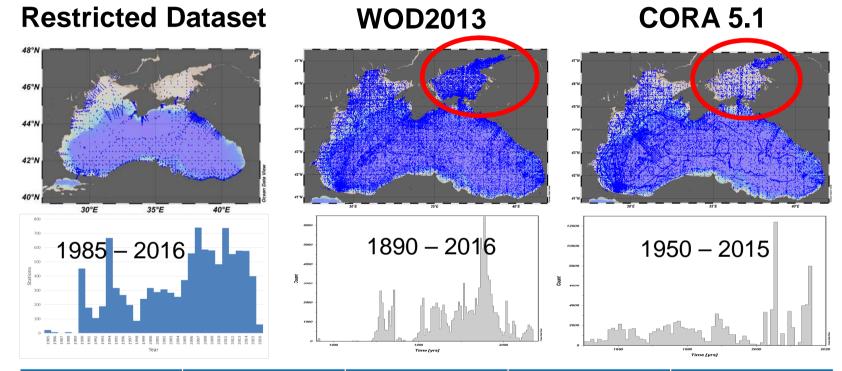
Work Plan and Timeline





Ongoing Activities: data integration Black Sea

Estimates from metadata: 60% duplicates 80000 stations (+50%)



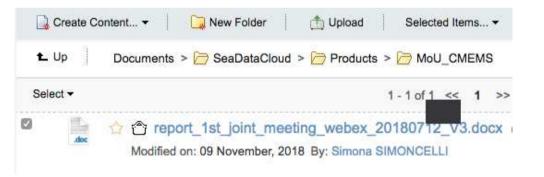
	unestricted	restricted	WOD2013	CORA 5.1
stations	137723	10528	120845	103721
	148	251		



Ongoing Activities: collaboration with CMEMS Ins TAC

1st Joint Meeting July 12th 2018





AGENDA

General introduction

- From SeaDataNet to SeaDataCloud: new data products and innovation
- 2. From phase I to phase II: CMEMS in situ TAC developments

DISCUSSION

Review of Quality Check procedures

- 3. SDN Quality Check Strategy: from visual inspection to automatic approac
- 4. CMEMS quality check procedures

DISCUSSION → find a common strategy

- 5. First release of SDC_xxx_DATA_TS_V1 and Product Information Documents (PIDocs)
- 6. CMEMS new data types?

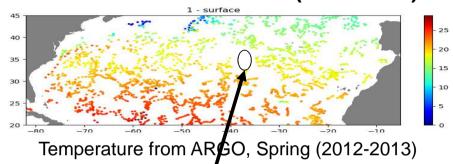
DISCUSSION --> data type approach/timelines/mutual requirements



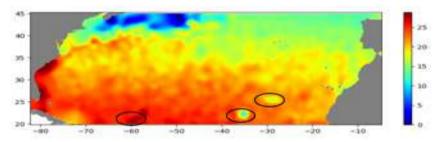
Advanced estimates of GLOBAL climatologies

K. Shahzadi, N.Pinardi, M.Zavatarelli (UNIBO), S.Lyubartsev (CMCC), S.Simoncelli (INGV)

World Ocean Database (WOD2013)



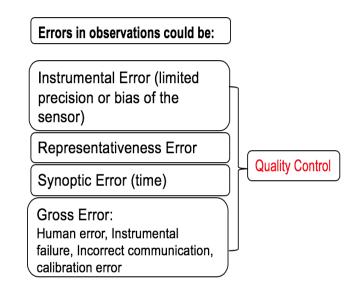
What's the **CLIMATOLOGICAL** value at this location?



DIVA estimate of 2012-2013 observations

Preliminary results from DIVA

- → anomalous features
- → better quality control to eliminate outliers



Need to develop a new NONLINEAR QC (NQC) to reduce the representatinevess error





Climatologies SDC_CLIM_TS_V1

- T and S monthly and seasonal climatologies with **increased horizontal** and **vertical resolution** (WOA standard depth) covering the time period 1955-2017
- Integration of SeaDataCloud data collections with external data sets to increase data coverage (CMEMS, WOD2018, ICES)
- Analysis of (1) space/time data distribution; (2) data types consistency; (3) long term variability to compute climatologies on a decadal basis (sliding decades when possible)
- Product validation → consistency analysis with WOA and CMEMS products (satellite reprocessed data sets and reanalysis)
- → PIDoc will contain all this information
- →SDC_CLIM_TS_V1 will be also accessible through EMODnet Physics



Conclusions

- Introduction of PIDocs represented a very good progress
- SDC_DATA_TS_V1 publication → data set paper and submission of SDN QCS as Ocean Best Practice
- Metadata analysis will be extended to all the regional seas in the next QCS loop
- Hunting of data omissions will be intensified





- QC analysis per data type will continue to assure data consistency
- Integration with external datasets (WOD2018, CMEMS) for climatology production (upcoming milestone)
- WP11 met regularly online to monitor and harmonize the activities
- High participation/great collaboration/very good progresses
- Deliverables and reports were all submitted in time



Open Issues

- underway data → data have been subsampled in SDC_DATA_TS_V1 release (1 over 7) → wish to manage underway data (FerryBox and TSG) separately and provide the full resolution (on going discussion)
- restricted data → need to reduce its percentage (unlock data older than 2010)
- Time series → fragmented and difficult to manage for QC and data products → how do we want to proceed? WP11 wish a reprocessing at the data center level
- Meeting with Ins-TAC CMEMS ASAP

Thanks to

Christine Coatanoan, IFREMER
Volodymyr Myroshnychenko, METU
Örjan Bäck, SMHI
Helge Sagen, IMR
Serge Scory, RBINS
Reiner Schlitzer, AWI
Michèle Fichaut, IFREMER
Dick Schaap, MARIS













