Product Information Document (PIDoc)

SeaDataCloud Temperature and Salinity Historical Data Collections for the North Sea (Version 1)

SDC_NS_DATA_DISCRETE_TS_V1
SDC_NS_DATA_TRAJECTORIES_TS_V1

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

Product Names

SDC_NS_DATA_DISCRETE_TS_V1

SDC_NS_DATA_TRAJECTORIES_TS_V1

Extended names

SeaDataCloud Temperature and Salinity Historical Data Collection for the North Sea, Discrete data (Version 1)

SeaDataCloud Temperature and Salinity Historical Data Collection for the North Sea, "Underway" data (Version 1)

Product DOIs

Serge Scory, Dick Schaap, Reiner Schlitzer (2018). North Sea - Temperature and salinity Historical Data Collection SeaDataCloud V1. http://doi.org/10.12770/d9310aa6-c6b3-4657-a3c9-b0a4dd0f016c

Short description

The SeaDataCloud Temperature and Salinity Historical Data Collection for the North Sea includes open access in situ data covering the period 1893 – 2017. The data were retrieved from the SeaDataNet infrastructure at the end of October 2017.

Authors	Contact
Serge Scory	serge.scory@naturalsciences.be

Dissemination	Copyright terms
Public	

How to Cite

History

Version	Authors	Date	Comments
V0	S. Scory	23/02/2018	First draft
V1	S. Simoncelli	23/02/2018	Revision
V2	S. Scory	27/02/2018	Revision

V3	S. Scory	25/04/2018	Update
V4	S. Simoncelli	20/06/2018	Revision
V5	S. Scory	27/09/2018	Update
V6	S. Simoncelli	12/10/2018	Final Check

Table of contents

Abstract

The SeaDataCloud Temperature and Salinity Historical Data Collection for the North Sea includes open access in situ data on temperature and salinity for the period 1893 – 2017. The data were retrieved from the SeaDataNet infrastructure at the end of October 2017. The dataset format is ODV binary collections [3]. The quality control of the data has been performed with the help of ODV software. Data Quality Flags have been revised following common recommended procedures defined under SeaDataNet 2 project [2] in conjunction with visual expert check. Trajectories ("underway" data) and "discrete" data (CTD, XBT, discrete water samplers) have been handled separately.

The final number of stations in the "discrete" collection is 162 462, compared to 580 376 in the "trajectories" collection.

Whenever SDC_NS_DATA_DISCRETE_TS_V1 product is used, this PIDoc should be cited in any publication. We also ask users to remember that hard-working scientists made these measurements, often under severe conditions. Further, the data providers normally possess insight on the quality and context of the data not always shared with the SeaDataCloud team. Hence, inviting data providers and product leaders to collaborate in scientific investigations that depend on their data and data products is considered good and fair practice. Importantly, this will promote further sharing of data and will be beneficial to science.

1. General description of the data collection

The SeaDataCloud Temperature and Salinity Historical Data Collection for the North Sea contains data on temperature and salinity of water body (profiles and surface measurements) retrieved from the SeaDataNet infrastructure in October 2017.The collection includes non-restricted data obtained from 85 organisations (originators and collating centres) listed in Annex 2 – Data providers.

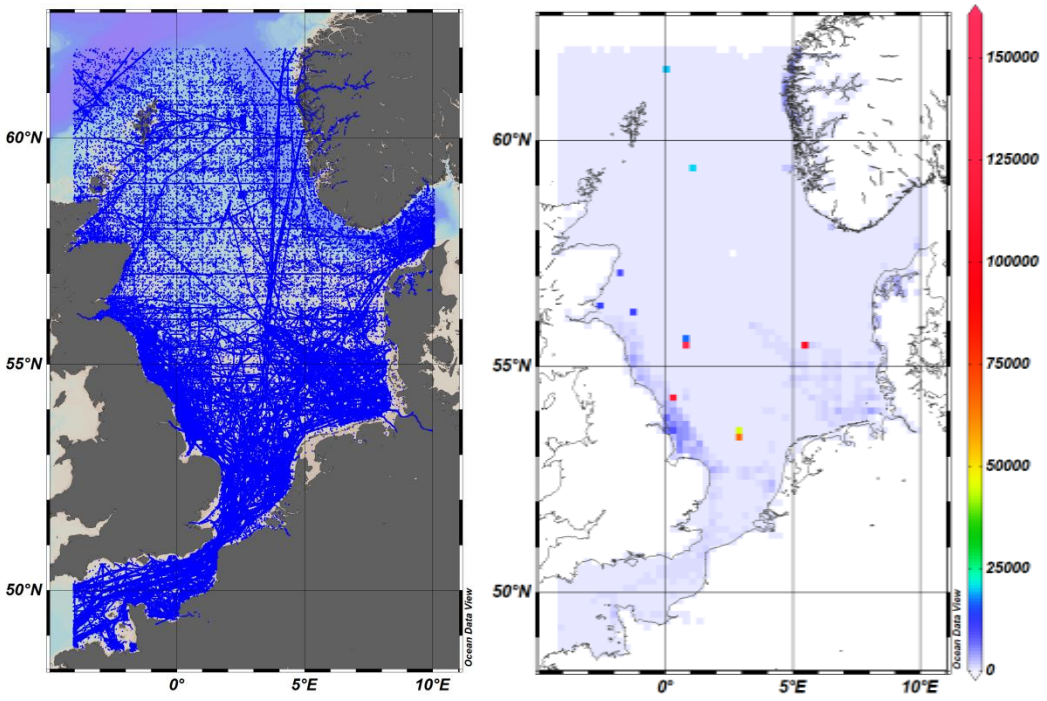
The collection covers the period 1893 – 2017. All data in the collection have been quality controlled according to procedures described in Annex 1. "Underway" data (FerryBox series, GOSUD series, *etc*.) were extracted and handled separately. Moorings have not been taken into consideration. Data statistics for the final collections are presented in Table 1.

Discrete	Stations	Values
Total	162 452	
Temperature	158 622	7 817 193
Salinity	157 545	7 707 384
T and S	153 880	7 699 641

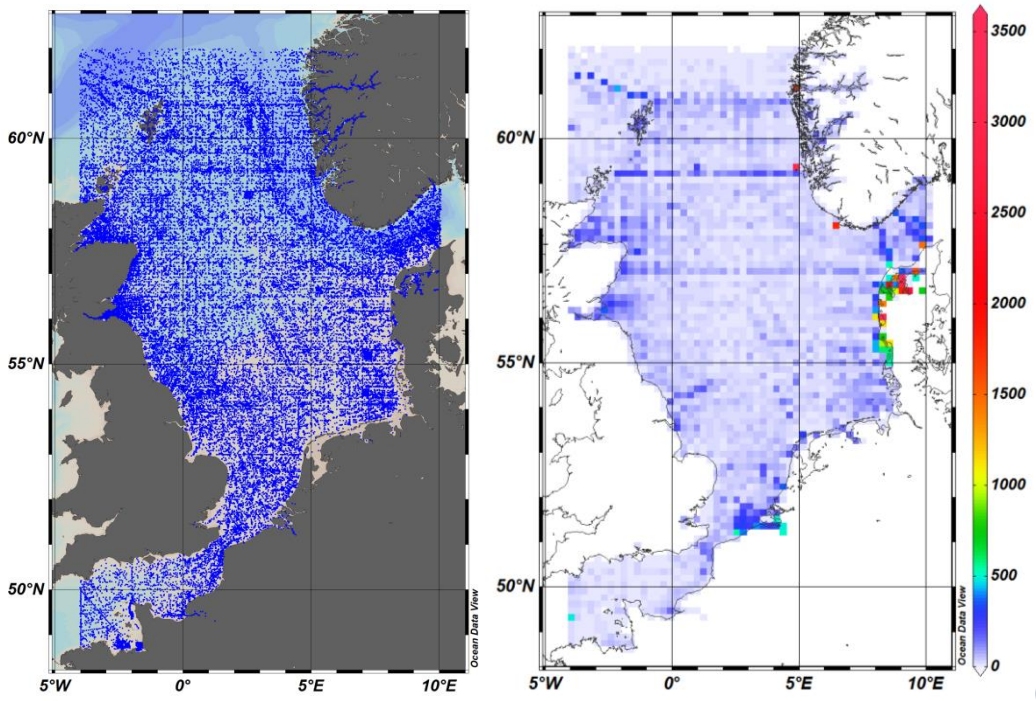
Table 1 Number of stations and of values in the final "discrete" data set.

Trajectories	Stations	Values
Total	580 376	
Temperature	576 356	576 356
Salinity	431 903	431 903
T and S	431 809	420 867

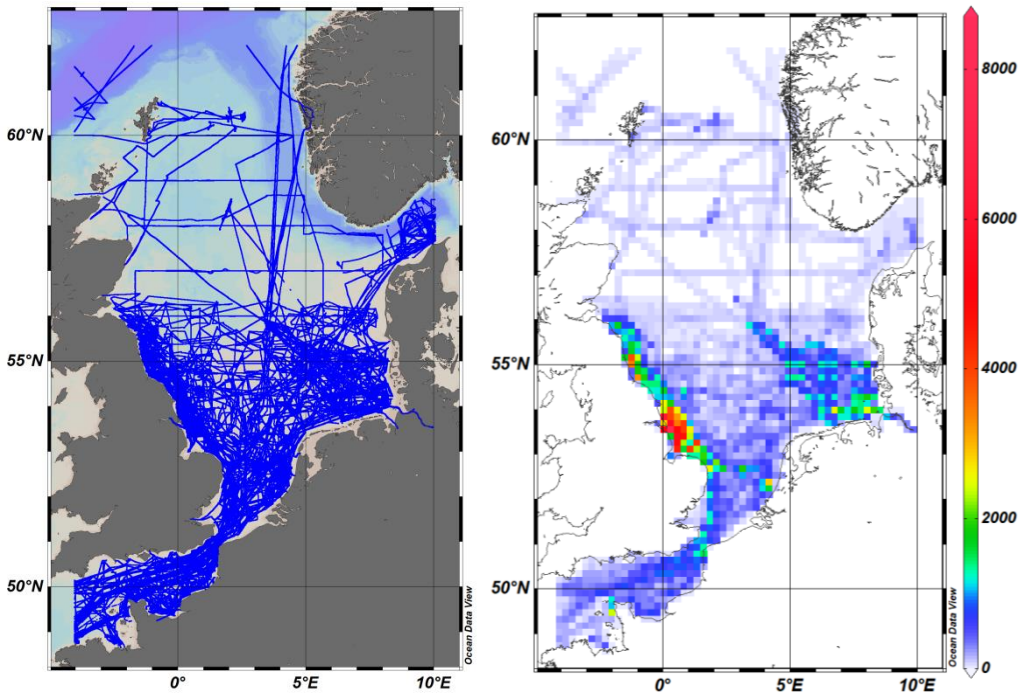
Table 2 Number of stations and of values in the final "trajectories" data set.


60°N
55°N
50°N
0°
5°E
10°E
150000
125000
100000
75000
50000
25000
0
Ocean Data View

(a)


60°N
55°N
50°N
5°W
0°
5°E
10°E
3500
3000
2500
2000
1500
1000
500
0
Ocean Data View

(b)

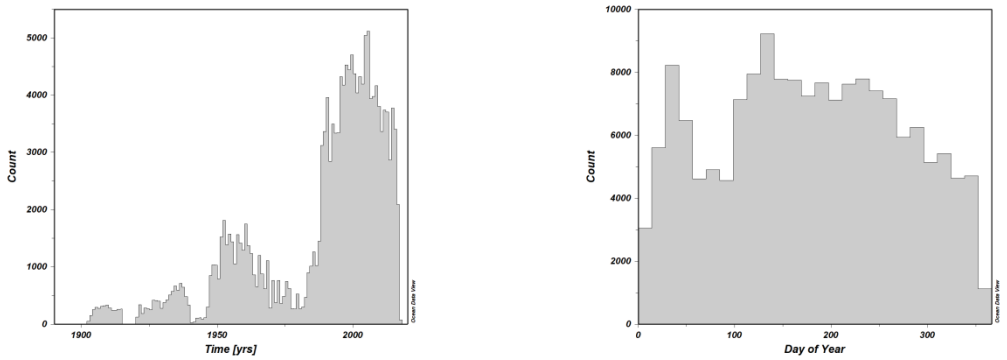


(c)

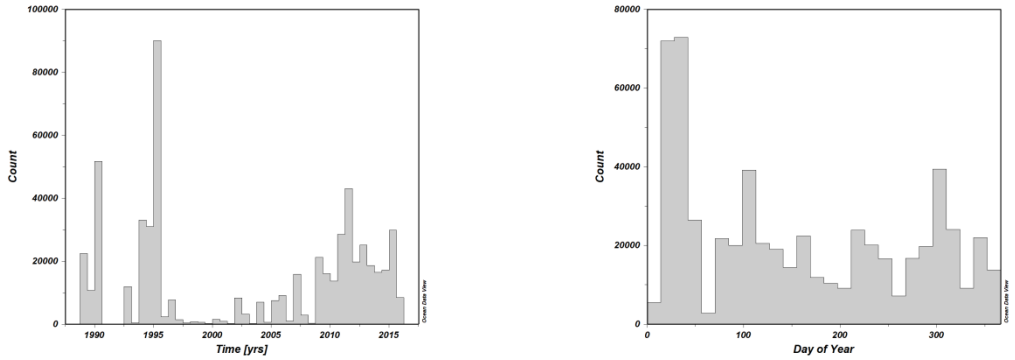
Figure 1 Spatial distribution of the stations (left) and their density (right). (a) Original collection; (b) Profiles and discrete samplers; (c) "Underway" data.

The spatial distribution original collection is dominated by data from fixed moorings that haven't been further taken into consideration (they should have been reported as time series). Both "discrete" and "underway" collections show high densities in certain areas, due to intensive monitoring or research programmes (Figure 1): Belgian coastal zone and Rhine/Meuse Delta, Danish coast, Skagerrak, Firth of Forth and two transects at the Northern boundary.

Although the collection spans from 1893 till 2017, most of the measurements were made during the last 30 years: 115415 discrete and profile stations (out of 162452) were performed after 1985 and underway measurements started in 1985. The coverage over the year is rather uniform (Figure 2).



(a)



(b)

Figure 2 Distribution of the measurements (stations) over time (left) and seasonal distribution (right). (a) Profiles and discrete samplers; (b) underway measurements.

2. Quality Control Procedure

The quality control work followed the best practices that were defined during the project SeaDataNet 2, see Annex 1 – QC Best Practices.

Errors found during the work in SeaDataNet 2 were checked in first instance as it was unsure whether our recommendations for improving the data had been implemented. Some systematic errors were still there but most of the time the overall quality had improved.

Using ODV, after handling obvious outliers in the whole data set we took two approaches: the first was to inspect the data by collating centre. Our experience has indeed shown that the routine applied at each centre for processing the data tends to generate systematic errors. The second approach was to work on sub-regions (Figure 3).

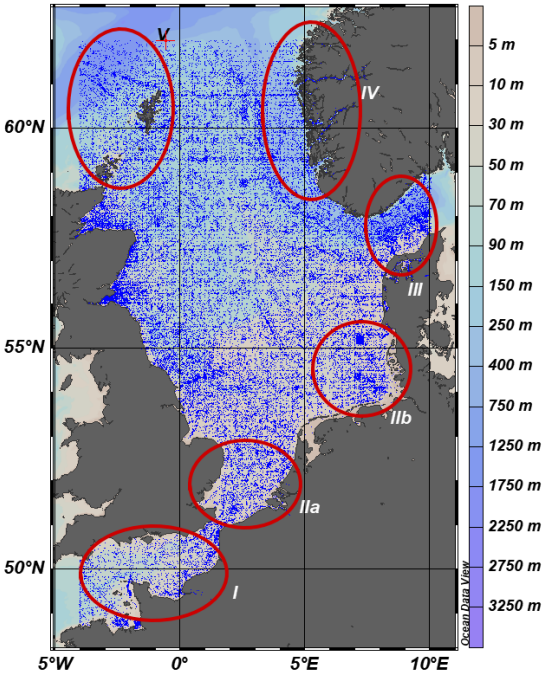


Figure 3 The various sub–regions used for the detailed quality control. I: Channel, IIa & IIb: shallow regions, respectively the Southern Bight and the German Bight; III: Skagerrak; IV: Norwegian fjords; V: Deep oceanic zone.

Whenever needed, data were checked "in context", *i.e.* by looking at all data on smaller geographical or temporal scale or for a given period of the year.

The DIVA analysis tool [6] incorporated in ODV has also been used for spotting anomalies.

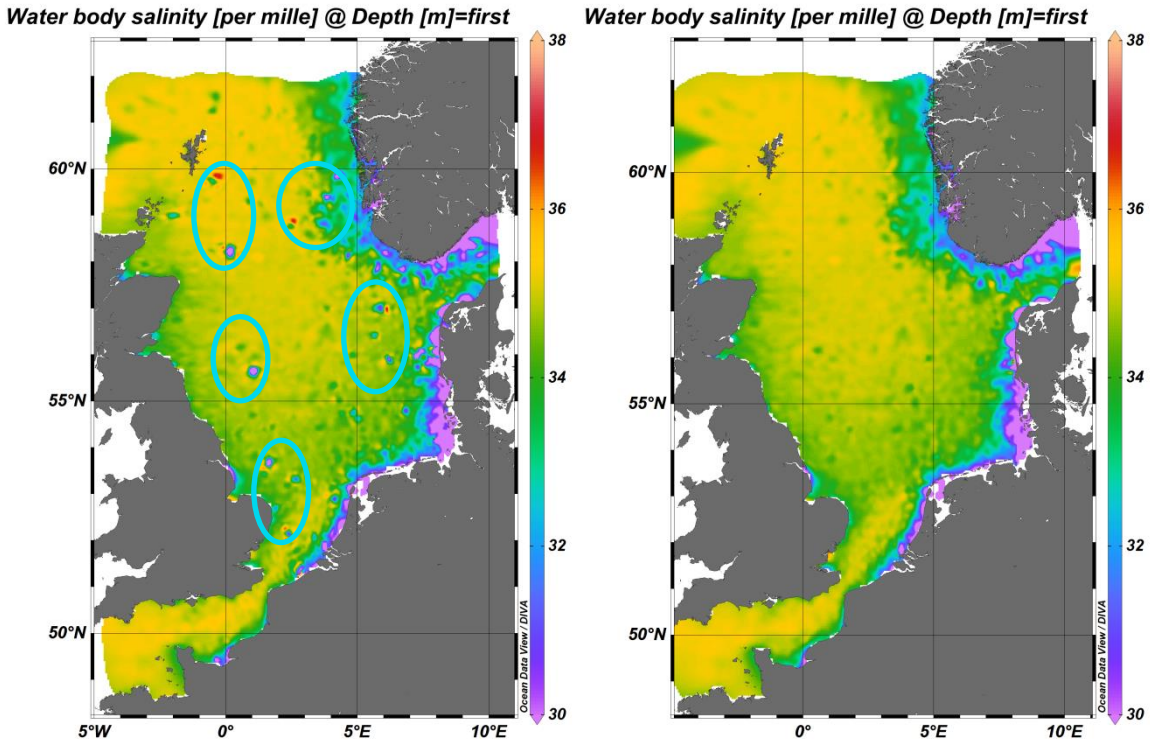


Figure 4 Examples of anomalies detected using DIVA in ODV. Maps show the Salinity field ("discrete" collection) in winter, before and after cleaning of the data.

3. Quality assessment results

All data have been quality controlled according to Section 2. *Quality Control Procedure*. At the end, around 19000 salinity values and 3200 temperature values are flagged as suspicious/bad, (flags 3 and 4). This is about 0.3% of the total amount of salinity values and 0.05% of the total amount of temperature values, further details can be seen below in Table 3.

The majority of the suspicious values come from profiles generated by instruments that were not stabilised at the beginning of the cast. This is mostly visible in profiles published by Belgian and Dutch data centres. There are two apparent causes: a common bad practice that produces (relatively) more bad results in shallow waters and the fact that the operators are interested in parameters that require less or no stabilisation for the instrument to provide rather accurate results.

Some data centres report both down-cast and up-cast in the same file. Data are not *per se* bad but as the data set only contains the depths of sampling and not the chronology, it is impossible to sort the data. In that case, we flagged the profile with “A”. For the same reason (data might be good), the two station on land were flagged with “A”.

A series of profiles in the Skagerrak consistently shows very high values of salinity. These values were also flagged as “A”. Other outliers were flagged 3.

Three sampling depths are negative. Final sampling depths were checked, when looking suspicious, against the latest (February 2018) data set published on the EMODnet Bathymetry portal, resulting in several of them being flagged as 3 or 4. As these profiles were all coming from the same centre we shall recommend them to check all their profiles; we indeed cannot check if there might be a systematic decoding error. (No conclusion can be drawn when the sampling depth is smaller than the local bottom depth.)

Compared to SDN_V2, the number of samples that have a flag “0” (i.e. not–QCed) inside a profile of which the other values are quality controlled has significantly decreased. But we sometimes found values flagged as “3” (probably bad) although they are perfectly in agreement with the neighbouring ones, which are flagged “1”.

A feature that would need to be further investigated is the impressive number of profiles with constant values along the depth. Although conditions of well mixing isn’t rare in large parts of the North Sea, perfectly constant values of salinity –for example–, up to the third decimal figure, over a depth of more than 100m, are counter–intuitive.

Quality flag	Number of S values	%
0	0	0
1	7 575 306	98.29
2	105 593	1.37
3	16 533	0.21
4	5 282	0.07
5	0	0
6	0	0
7	0	0
8	4 245	0.06
9	0	0
A	0	0
Total	7 706 959	

Quality flag	Number of T values	%
0	0	0
1	7 722 333	98.79
2	90 972	1.16
3	1 340	0.02
4	813	0.01
5	0	0
6	0	0
7	0	0
8	1 209	0.02
9	0	0
A	0	0
Total	7 816 667	

Table 3 "Discrete" collection: Number of S & T values for each quality flag, where the QF for depth parameter is equal to 1 (good) or 2 (probably good) after quality control of the data set.

Quality flag	Number of S values	%
0	0	0
1	419 787	97.19
2	1 082	0.25
3	10 453	2.42
4	581	0.13
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
A	0	0
Total	431903	

Quality flag	Number of T values	%
0	0	0
1	564 124	97.88
2	128	0.02
3	11 028	1.91
4	1 076	0.19
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
A	0	0
Total	576 354	

Table 4 "Trajectories" collection: Number of S & T values for each quality flag, where the QF for depth parameter is equal to 1 (good) or 2 (probably good) after quality control of the data set.

The following Figures show the "discrete" data set scatter plots before and after cleaning: Temperature over depth (Figure 5), Salinity over depth (Figure 6) and TS diagram (Figure 7 and Figure **8**).

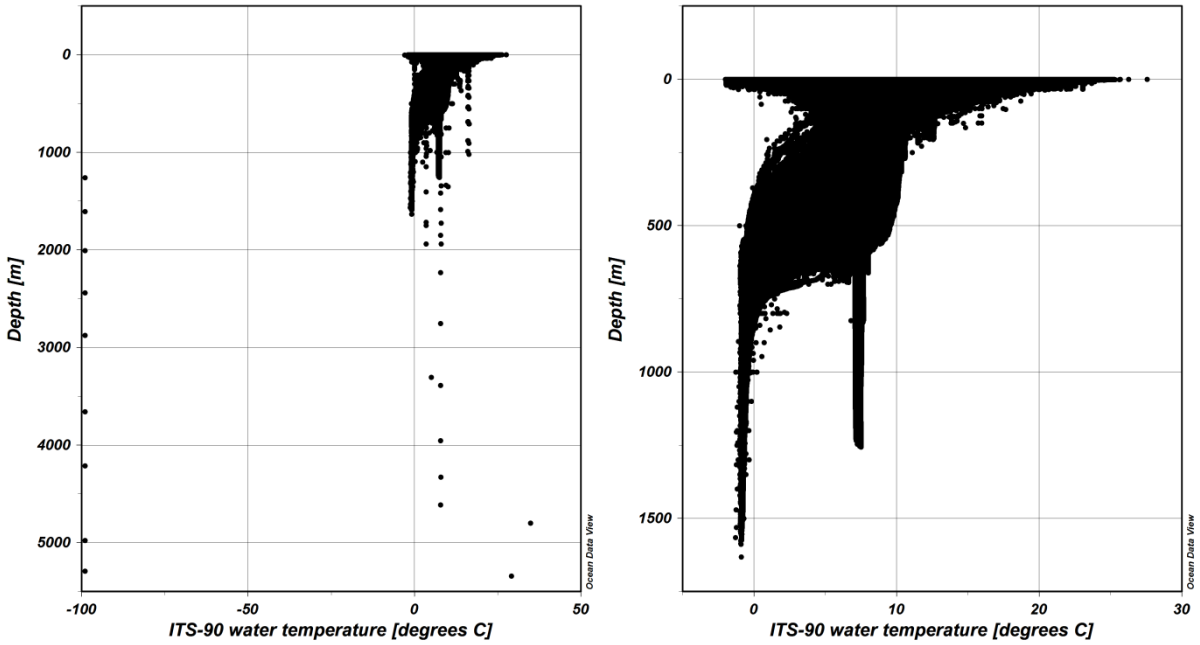


Figure 5 Temperature: data with quality flags set to 0 ("not controlled"), 1 ("good"), 2 ("probably good"), 5 ("changed value") and 8 ("interpolated value") in the original data set (left) and after quality control (right).

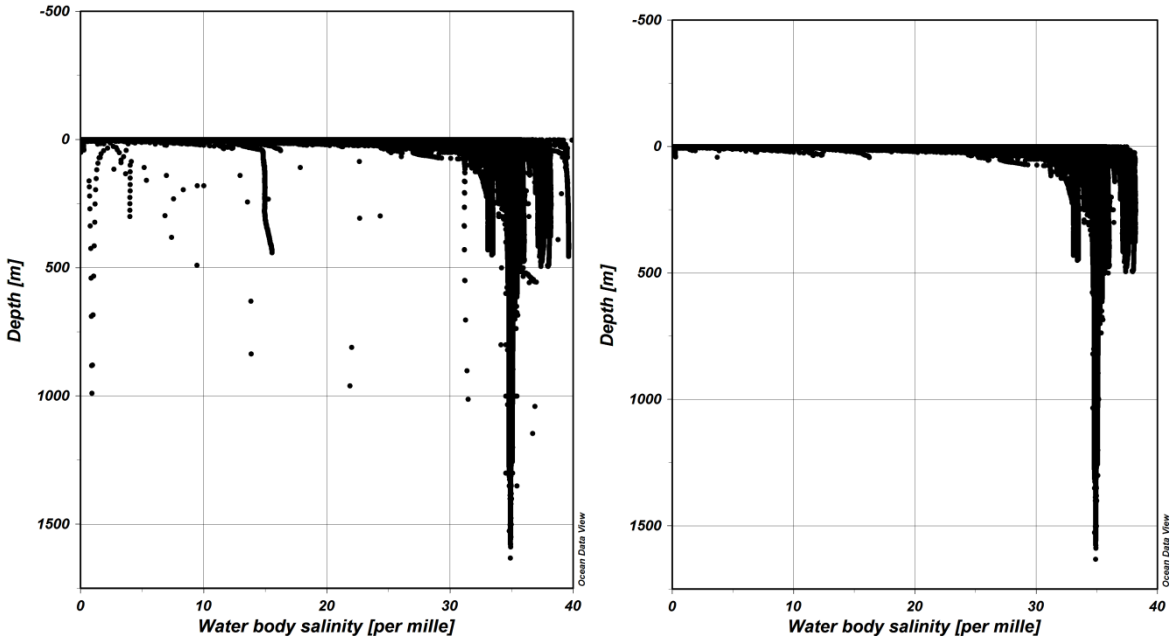


Figure 6 Salinity: data with quality flags set to 0 ("not controlled"), 1 ("good"), 2 ("probably good"), 5 ("changed value") and 8 ("interpolated value") in the original data set (clipped to the 0–40 range, left) and after quality control (right).

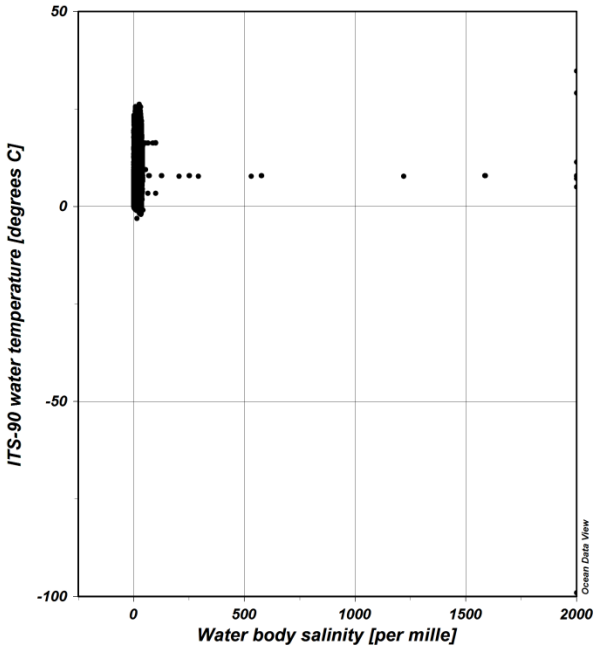


Figure 7 TS-diagram of the original data set.

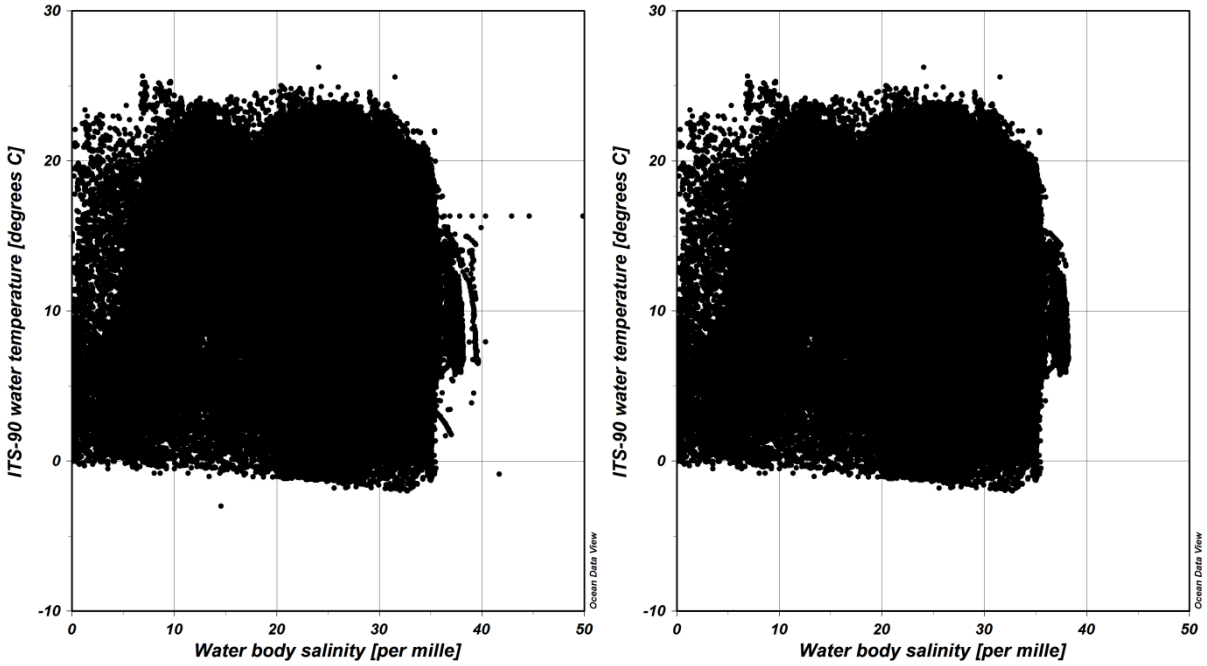


Figure 8 **TS-diagram of the original data set (zoomed, left) and of the data set after quality control (right).**

4. Technical Specifications

4.1. “Discrete” collection

Product Format:

Ocean Data View (ODV) collection.

Data Policy:

No limitations on usage, however for data access the registration is required at http://www.marine-id.org.

Product Usability:

The collection contains unique large array of quality controlled data on temperature and salinity of North Sea waters for the period 1893 – 2017 that can be used to support general oceanographic studies, such as investigation of variability of physical properties or long time trends, as well as applications such as circulation models. The data contained in the collection were measured using profiling devices (XBT, CTD…) or discrete water samplers (bottles).

Whenever SDC_NS_DATA_DISCRETE_TS_V1 product is used, this PIDoc should be cited in any publication. We also ask users to remember that hard-working scientists made these measurements, often under severe conditions. Further, the data providers normally possess insight on the quality and context of the data not always shared with the SeaDataCloud team. Hence, inviting data providers and product leaders to collaborate in scientific investigations that depend on their data and data products is considered good and fair practice. Importantly, this will promote further sharing of data and will be beneficial to science.

Changes since previous version:

The previous version of the product was released in framework of the SeaDataNet 2 project and available in SEXTANT product catalogue [4] under the name “North Sea - Temperature and salinity observation collection V2” (SDN_V2) [5]. The spatial domain is the same but its time coverage extended from 1900 – 2014 to 1893 – 2017.

At the time of creating the SDN_V2 collection we didn’t handle the high frequency series separately. Therefore this collection contains a very large number of stations and values, as shown in Table 5.

For the sake of the comparison, we looked at a subset of SDN2_V2, containing discrete data only. The statistics of this subset are compared to those of SDC_V1 in Table 6. These figures show a global increase of the number of station of more than 40%, resulting in an increase of 17% for the temperature data and of 19% for the salinity data and for the T,S-pairs.

	SDN2_V2 #stations	% of total	#values
Total	1 610 854		
Temperature	1 590 331	98.73%	15 943 941
Salinity	946 724	58.77%	7 287 495
T and S	944 129	58.61%	7 344 660

Table 5 Number of stations and values in in the full SDN_V2 collection and of the number of those containing respectively T values, S values and T,S-pairs.

	Subset (profiles) of SDN2_V2 #stations	% of total	#values	SDC_V1 #stations	% of total	#values
Total	115 596			162 452		
Temperature	112 422	97.25%	6 670 529	158 622	97.64%	7 817 193
Salinity	110 719	95.78%	6 473 381	157 545	96.98%	7 707 384
T and S	108 272	93.66%	6 473 721	153 880	94.72%	7 699 641

Table 6 Comparison of the number of stations and values in the subset of SDN_V2 containing only discrete data and in SDC_V1, and of the number of those containing respectively T values, S values and T,S-pairs.

4.2. "Trajectories" collection

Product Format:

Ocean Data View (ODV) collection.

Data Policy:

No limitations on usage, however for data access the registration is required at http://www.marine-id.org.

Product Usability:

The collection contains unique large array of quality controlled data on temperature and salinity of North Sea waters for the period 1988 – 2017 that can be used to support general oceanographic studies, such as investigation of variability of physical properties or long time trends, as well as applications such as circulation models. The data contained in the collection "underway" data (*e.g.* FerryBox).

Whenever SDC_NS_DATA_TRAJECTORIES_TS_V1 product is used, this PIDoc should be cited in any publication. We also ask users to remember that hard-working scientists made these measurements, often under severe conditions. Further, the data providers normally possess insight on the quality and context of the data not always shared with the SeaDataCloud team. Hence, inviting data providers and product leaders to collaborate in scientific investigations that depend on their data and data products is considered good and fair practice. Importantly, this will promote further sharing of data and will be beneficial to science.

Changes since previous version:

N/A.

Annex 1 – QC Best Practices

- The basic QC analysis steps applied during SeaDataNet2 Project using ODV were:
- **Data coverage**;
- **Data distribution maps** per Temperature, Salinity and TS couples;
- **Data density maps** (domain binning);
- **Time coverage and time distribution** → histograms with annual, seasonal and monthly data distribution;
- TS scatter plots of the entire dataset;
- Scatter plot of observations with QF=1 (good) and QF=2 (probably good);
- Scatter plot observations with QF=0 (no quality check);
- **Gross range check** to detect observations with temperature and salinity out of reasonable values;
- Visual control of scatter-plots to identify wrong profiles (outliers);
- Identification of stations falling on land;
- Identification of stations having unreal depth;
- Identification of wrong or missing data;
- Stability check on density

Additional checks are advisable per specific:

- **areas** with similar hydrodynamic characteristics;
- **layers** (surface, intermediate, bottom);
- **time periods** (decades, or specific periods i.e. Eastern Mediterranean Transient, Western Mediterranean Transition, Norther Ionian Reversal);
- **Instrument type** → consistency issue of historical data;

Duplicate Check is another important step when performing SDC data integration with external data sources for climatologies and new data products generation.

Annex 2 – Data providers

Collating centres are typeset in **bold**.

EDMO	Originator/Collating centre
13	University of Plymouth, Institute of Marine Studies
17	National Oceanography Centre, Southampton
20	University of Wales, School of Ocean Sciences
38	Fisheries Research Services, Aberdeen Marine Laboratory
43	**British Oceanographic Data Centre**
44	Scottish Association for Marine Science
47	Plymouth Marine Laboratory
90	Environment Agency Head Office
96	**German Oceanographic Datacentre (NODC) (DOD)**
240	Université Libre de Bruxelles, Laboratory of Chemical Oceanography and Water Geochemistry
396	**Marine Institute**
422	**Flanders Marine Institute**
429	Royal Belgian Institute of Natural Sciences, Management Unit of North Sea and Scheldt Estuary Mathematical Models, Ostend
431	Royal Belgian Institute of Natural Sciences, Management Unit of North Sea and Scheldt Estuary Mathematical Models
443	Université Libre de Bruxelles, Ecology of Aquatic systems
486	**IFREMER / IDM / SISMER - Scientific Information Systems for the SEA (486)**
509	IFREMER / STATION PORT EN BESSIN
521	Roscoff Marine Station, Sorbonne University, UPMC Univ Paris 06 and CNRS
527	IFREMER / RBE Department / Biogeochimical end Ecotoxicological Resarch Unit (Nantes)
540	**Shom**
542	IFREMER / HMMN-DEPARTEMENT HALIEUTIQUE DE MANCHE-MER DU NORD
545	**Swedish Meteorological and Hydrological Institute**
612	**Institute of Marine Research - Norwegian Marine Data Centre (NMD)**
630	**NIOZ Royal Netherlands Institute for Sea Research**
681	**All-Russia Research Institute of Hydrometeorological Information - World Data Centre (RIHMI-WDC) National Oceanographic Data Centre (NODC)**
682	Atlantic Scientific Research Institute for Marine Fishery and Oceanography
713	**Marine Systems Institute at Tallinn University of Technology**
727	**Marine Hydrophysical Institute**
729	**Aarhus University, Department of Bioscience, Marine Ecology Roskilde**
730	**International Council for the Exploration of the Sea (ICES)**
756	Far Eastern Regional Hydrometeorological Research Institute
894	Lithuanian Hydrometeorological Service under the Ministry of Environment
903	Murmansk Hydrometeorological Administration of Roshydromet
907	Navy Main Administration of Navigation and Oceanography, Ministry of Defence

919	Polar Scientific Research Institute of Fishery and Oceanography
931	Odessa Branch of SOI (State Oceanographic Institute)
989	Federal Research Centre for Fisheries (Cuxhaven)
990	Federal Research Centre for Fisheries (Hamburg)
1016	IFREMER / DYNECO- Coastal Environment Dynamics department
1019	Wimereux Marine Station (University of Sciences and Technologies of Lille 1)
1031	IFREMER / CENTRE MANCHE - MER DU NORD
1039	CEA / INSTITUT DE RADIOPROTECTION ET DE SURETE NUCLEAIRE
1051	UNKNOWN
1054	IFREMER
1102	Danish Institute for Fisheries Research, Dept. of Marine Fisheries
1181	State Agency for Nature and Environment of Schleswig Holstein (LANU)
1181	State Agency for Nature and Environment of Schleswig Holstein (LANU)
1326	Institute of Biochemistry and Food Chemistry, University Hamburg
1327	Lower Saxony Water Management, Coastal Defense and Nature Conservation Agency
1351	Institute of Marine Research
1368	Alfred-Wegener-Institute for Polar- and Marine Research
1477	STATION MARINE DE DINARD
1526	Rijkswaterstaat Water, Traffic and Environment
1570	Thünen-Institute of Sea Fisheries (TI-SF)
1571	Thünen-Institute of Fisheries Ecology (TI-FI)
1575	Federal Research Institute for Rural Areas, Forestry and Fisheries (VTI)
1576	Institute of Fisheries Ecology - Cuxhaven (VTI-CUX)
1578	**Management Unit of North Sea and Scheldt Estuary Mathematical Models, Belgian Marine Data Centre**
1780	Faroese Fisheries Laboratory
1803	University of Southampton Department of Oceanography
1804	University of Southampton Department of Oceanography
1850	**Federal Maritime and Hydrographic Agency**
1919	Universite du Littoral Cote d'Opale / Oceanology and Geosciences Laboratory - UMR 8187 (LOG)
2090	Scottish Office Agriculture Environment and Fisheries Department - Aberdeen Marine Laboratory
2091	Scottish Office Agriculture and Fisheries Department - Aberdeen Marine Laboratory
2092	Department of Agriculture and Fisheries for Scotland - Aberdeen Marine Laboratory
2118	Institute of Oceanographic Sciences Wormley Laboratory
2133	Proudman Oceanographic Laboratory
2134	Institute of Oceanographic Sciences, Bidston Laboratory
2135	Marine Scotland Science
2195	DTU Aqua - National Institute of Aquatic Resources, Technical University of Denmark

2303	Department of Safety and Quality of Milk and Fish Products (Max Rubner)
2415	Polytechnic South West Institute of Marine Studies
2424	National Oceanography Centre, Liverpool
2492	Newcastle University Department of Marine Science and Coastal Management
2529	Dunstaffnage Marine Laboratory
2530	Scottish Marine Biological Association
2537	State Office for Agriculture, Environment and Rural Areas of Schleswig Holstein (LLUR)
2947	GEOMAR Helmholtz Centre for Ocean Research Kiel
3272	Scottish Environment Protection Agency, Edinburgh Office
3326	IHE, Management Unit of North Sea and Scheldt Estuary Mathematical Models, Ostend
4547	Environment Agency South West Regional Office
4551	Environment Agency South East Regional Office
4552	Environment Agency Solent and South Downs Area Office
4554	Scottish Environment Protection Agency, Angus Smith Building

References

1. SeaDataCloud project (2016-2020), grant agreement 730960, EU H2020 programme, www.seadatanet.org/About-us/SeaDataCloud
2. Seadatanet 2 project (2011-2015), grant agreement 283607, EU Seventh Framework Programme, https://www.seadatanet.org/About-us/SeaDataNet-2
3. Schlitzer, R., Ocean Data View, odv.awi.de, 2017
4. Sextant catalogue, Seadatanet products, http://sextant.ifremer.fr/en/web/seadatanet
5. Digital Object Identifier (DOI, https://www.doi.org/) for the Seadatanet 2 historical data collection V2, http://dx.doi.org/10.12770/8a51f275-6a8a-4ac2-ba7e-fe491e63a17d
6. Troupin, C., Barth, A., Sirjacobs, D., Ouberdous, M., Brankart, J.-M., Brasseur, P., Rixen, M., Alvera Azcarate, A., Belounis, M., Capet, A., Lenartz, F., Toussaint, M.-E., & Beckers, J.-M. (2012). Generation of analysis and consistent error fields using the Data Interpolating Variational Analysis (Diva). Ocean Modelling, 52-53, 90-101.

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
NS	North Sea
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