SeaDataNet II data products: the North Atlantic Ocean Region

Christine Coatanoan, Simona Simoncelli, Reiner Schlitzer, Michèle Fichaut, Dick Schaap, Sissy Iona
IMDIS, Gdansk, October 2016
SeaDataNet II (SDN) EU-project

- Implementation of a quality control (QC) strategy, continuously reviewed to improve the quality of the global dataset for creation of the best products.
- Strategy developed in collaboration with MyOcean In-Situ Thematic Assemble Centre (INS-TAC) at regional levels
- Temperature and salinity historical data collections were created by sea basins, as aggregated datasets and climatology products, and covering the time period 1900-2013.
• **Specific QC procedure**

  ➢ Implementation of a specific procedure to assure and certify the best quality for the datasets:

  • After the dataset harvesting from the central CDI catalogue, QC has been performed at regional levels in a coordinate way, using the ODV software as a common and basic QC analysis tool.

  • Those datasets have also been scrutinized by the MyOcean regional coordinators, which have sent feedbacks to the SDN regional partners.
Aggregation dataset – different steps

CDI extraction of all the datasets
SeaDataNet Licence
(CDI: Discovery and data access)

To deliver to MyOcean In situ TAC:
the raw aggregated data set for the time period 1990-2012
a list of outliers with the ranges defined for depth, temperature and salinity;
a report which briefly describes the data collection (maps, histograms, scatter plots)

To send to the NODCs:
a list of data with QC=0 (no QC analysis performed) for the entire time period 1900-2012
a list of outliers for the entire time period 1900-2012
a report with the general description of the entire data collection (1900-2012)
Quality Check Strategy implemented during SDN project

This loop allowed to highlight doubtful data and to organize the data anomalies in lists that have been sent to each concerned data originator together with a guideline to explain the expected corrections. This implemented QC strategy involved the National Oceanographic Data Centers (NODC), on the base of those lists, to check and eventually correct the original data and then to resubmit the corrected data in the SDN dataflow. The QC procedure has also been designed to be iterative in order to facilitate the update and improvement of SDN database content.
Detailed descriptions for the North Atlantic Ocean

- General description of the dataset,
- Data quality assessment procedure and results

During SDN, several releases have been produced and the insertion of new data has showed a large increase of the data collection for the North Atlantic Area. Regarding the number of stations, only a small number of data have been detected as bad.
Increase of the dataset

<table>
<thead>
<tr>
<th>PAR</th>
<th># stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAR</td>
<td>1807266</td>
</tr>
<tr>
<td>T</td>
<td>1693840</td>
</tr>
<tr>
<td>S</td>
<td>785476</td>
</tr>
<tr>
<td>TS</td>
<td>784015</td>
</tr>
</tbody>
</table>

Dataset (2013) V1

V1 -> V1.1

N=431974

Dataset (2014) V1.1

N=1049547

~72% more/V1.1

N=1807266

Dataset (2015) V2

V2
There are still some anomalies. Ex of anomalies: inversion in depth (deeper=surface)??
Ex of anomalies: very huge bad value with QC 1 ???
First run and first bad data
### In terms of %

<table>
<thead>
<tr>
<th></th>
<th>TOT</th>
<th>QF0</th>
<th>QF1</th>
<th>QF2</th>
<th>QF:3-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>73877673</td>
<td>1200693</td>
<td>72438070</td>
<td>374</td>
<td>238536</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1,62%</td>
<td>98,05%</td>
<td>0,0005%</td>
<td>0,33%</td>
</tr>
<tr>
<td></td>
<td>T after correction</td>
<td>73877673</td>
<td>988646</td>
<td>72109310</td>
<td>374</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1,34%</td>
<td>97,6%</td>
<td>0,0005%</td>
<td>1,05%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>TOT</th>
<th>QF0</th>
<th>QF1</th>
<th>QF2</th>
<th>QF:3-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>34036664</td>
<td>1301510</td>
<td>32244014</td>
<td>46811</td>
<td>444329</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3,82%</td>
<td>94,73%</td>
<td>0,13%</td>
<td>1,30%</td>
</tr>
<tr>
<td></td>
<td>S after correction</td>
<td>34036664</td>
<td>1301510</td>
<td>31814244</td>
<td>46811</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3,82%</td>
<td>93,47%</td>
<td>0,13%</td>
<td>2,57%</td>
</tr>
</tbody>
</table>
Climatology - DIVA

- Diva 4.6.9
- Reference field (semi-normed analysis)
- Error field defined as “clever mean error field” (ispec=111)

- Variable: Temperature & Salinity
- Year: 1900 to 2013
- Monthly: 01 to 12
- Season: 1202 – 0305 – 0608 – 0911
- Depth: (IODE) 33 levels - 0 to 5500
  - used 0 to 4000
    - (5500, 5000, 4500, 4000, 3500, 3000, 2500, 2000, 1750, 1500, 1400, 1300, 1200, 1100, 1000, 900, 800, 700, 600, 500, 400, 300, 250, 200, 150, 125, 100, 75, 50, 30, 20, 10, 0)
- Data: public and restricted

Time-space resolution

SeaDataNet2 DIVA Contour & Mask

Contour levels

@seadatar
Salinity maximum at 1000m showing the Mediterranean outflow.
DIVA (error masking - 30% threshold)
Validation


Fig. 15.4. Temperature (°C) (a) and salinity (b) in the North Atlantic Ocean at 1000 m depth.

SeaDataNet2 DIVA Analysis - Temperature.19002013.0707
SeaDataNet2 DIVA Analysis - Salinity.19002013.0707

SeaDataNet2 DIVA Analysis - Temperature.19002013.0707
SeaDataNet2 DIVA Analysis - Salinity.19002013.0707

Map of annual-mean temperature at 100m below the surface in the NW Atlantic with a schematic representation of circulation

Distributions of salinity at 100 m below the surface in the northern North Atlantic.

Red arrows indicate the Gulf Stream and associated North Atlantic Current, which transport warm saline surface water.

Blue arrows indicate the East and West Greenland and Labrador Currents, which carry relatively cold and fresh water southward.

Error field
need more data
Conclusions

- QC strategy allowed to improve the QC on some NODCs.
- First validations of the products showed very specific structure in comparison with scientific literature.
- The final dataset aimed to study in details the water masses circulation in this area, but integration of new data should improve the quality of the product.
- ODV qualified dataset collections: available through the SDN web catalog at http://sextant.ifremer.fr/en/web/seadatanet/
- Future releases (SeaDataCloud) should have to more sustain the QC strategy and encourage NODCs to provide new data and take into account the data quality assessment outcomes. (already seen from the end of SDN II: data submission always increasing)
Thank you!!