# SeaDataNet

# WP10 Aggregated data set The Baltic Sea

- Compare v1 (2013) and v1.1 (2014)
- QC routine
- QC results

Örjan Bäck, Magnus Wenzer SMHI





## Amount of data in V1 and V1.1



SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

# V1.1 compared to V1

- ~34000 more CDIs
- Almost 3 000 000 more data points
- A couple of bugs in ODV detected and corrected in V1.1
- Merging of 2 or more profiles into one
- Duplication of temperature from one depth to two, creating a "staircase" profile



# Previous version, including bugs

### 3 profiles merged into one

(seemed to happen for profiles close in time and position, not very common)

#### 0 0 10 10 20 20 Pressure [dbar] Pressure [dbar] 30 30 40 40 Data View Ocean Data View 50 50 Ocean 5 0 2 3 6 2 3 5 6 1 Salinity [psu] Salinity [psu]

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

Current version, bugs fixed 3 profiles showing up correct



Previous version, including bugs. Temperature from first depth appeared on both first and second depth. Temperature from second depth appeared on third and fourth depth, and so on... "Staircase bug"

(a test checking 604 profiles showed that this bug appeared on 75 of those profiles, possibly up to 10% of the data)



Current version, bugs fixed Temperature profile shows correct data

(Same data in both plots)



# SeaDataNet Aggregated data set v1.1





### **Seasonal distribution**



### Annual distribution



Very high numbers are from ferry box trajectories, which contain many data points from only one depth and with a short timespan, typically a week.

Geographical distribution map also shows the dominating ferry box data.



## **QC** Salinity

- Salinity ~ 0-36, large geographic variation -> QC on subregions
- Even with subregions there are large amounts of data which makes visual inspection very hard. To handle this I looked at parts of the data, -1978 and after that all data for one year.
- Checking the density depth derivative for stable profiles was difficult because of the different depth resolutions, ranging from ~10 centimeters up to several meters.
- Solution: scroll through all profiles to visually inspect, to detect spikes and unstable density profiles.





### **QC** Temperature

- Temperature, large seasonal variation, from below 0°C in the winter to above 20°C in the summer, ~ -2 – 25 -> "monthly"QC
- Data was checked for 1-2 month(s) at a time, and because of the large data amounts we used even smaller subregions, se example below.
- Visual inspection by scrolling through the subset of profiles in ODV to detect spikes and outliers(and unstable density profiles).



January-February





# **QC** Results

- Depth (some bad whole profiles that were too deep, and some negative values)
- Salinity (mostly spikes which caused unstable density profiles (density not increasing with depth))
- Temperature (spikes/outliers)
- Position, 1-2 suspicous/bad positions
- ~1600 unique CDIs with suspicous and/or bad data, < 1% (CDI means in this case a unique data file corresponding to only one CDI where 1 or more suspicious values were flagged)</li>
- Flagged as 3 or 4 (probably bad or bad) and excluded from MyOcean data delivery and from DIVA

# **Qc** restricted data

18 CDIs contain suspicous/bad data (spikes), ~0.2%

### To do: compile the QC results and send to relevant SDN partners.



### **Before and after QC**





- Some DIVA tests have been done
- Still need to tune DIVA parameters
- and create seasonal background fields

