



Preliminary Analysis of the Black Sea raw aggregated T-S dataset

Volodymyr Myroshychenko

Devrim Tezcan

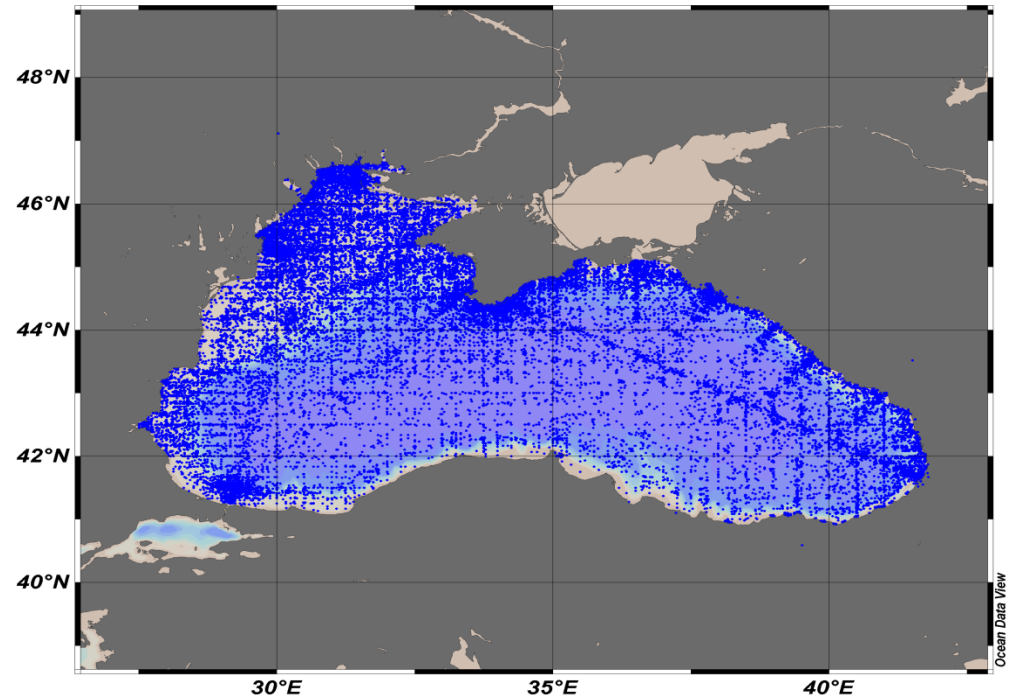
IMS METU

SeadataNet Regional Coordinators Webex
Conference

4/10/13

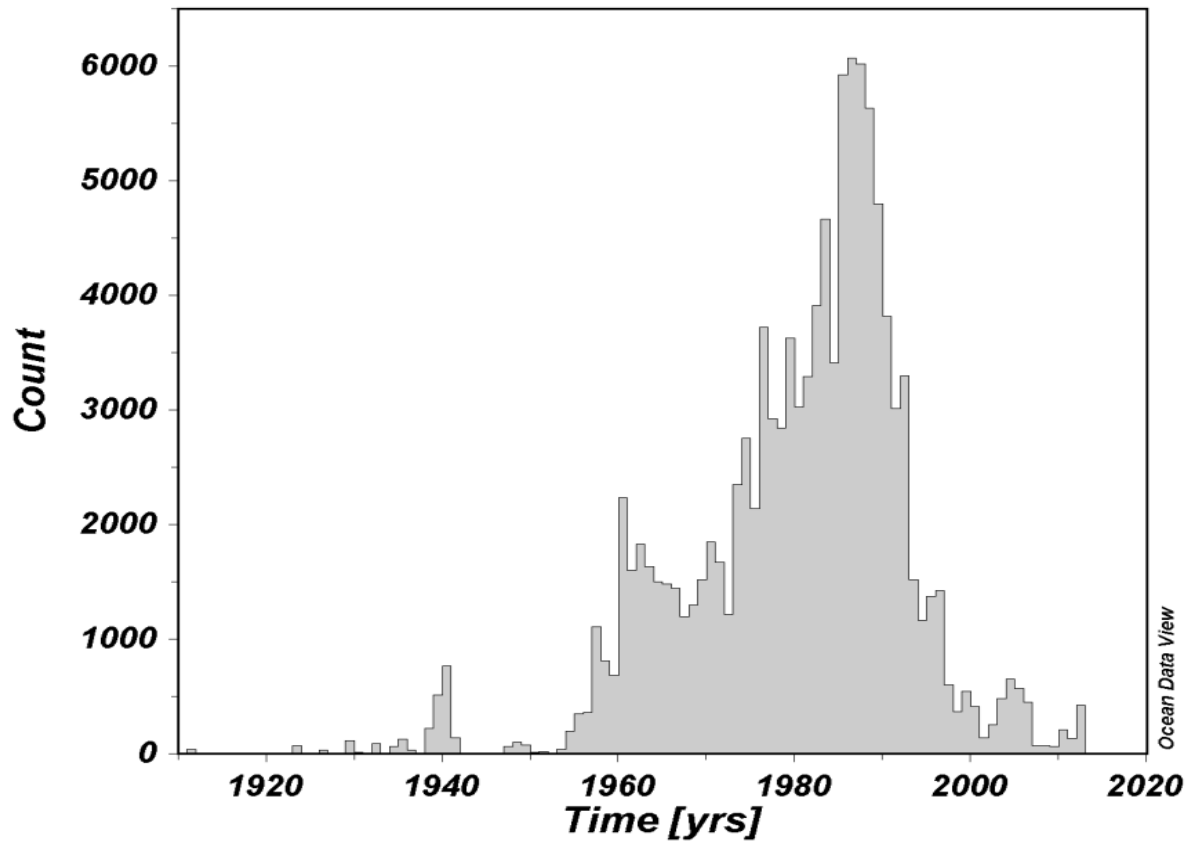
General statistics*

- 2037 cruises
- 114694 stations
- 3465204 samples
- Period: 1900-2012

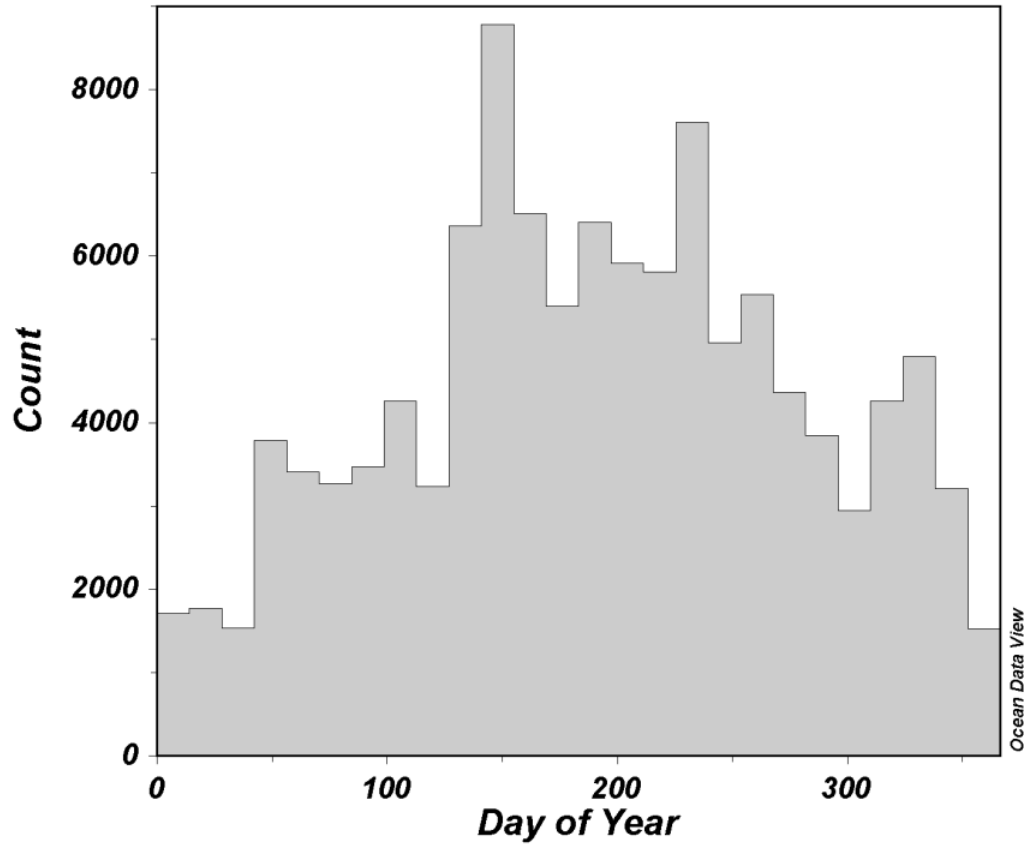


* After excluding the Sea of Marmara and Sea of Azov data

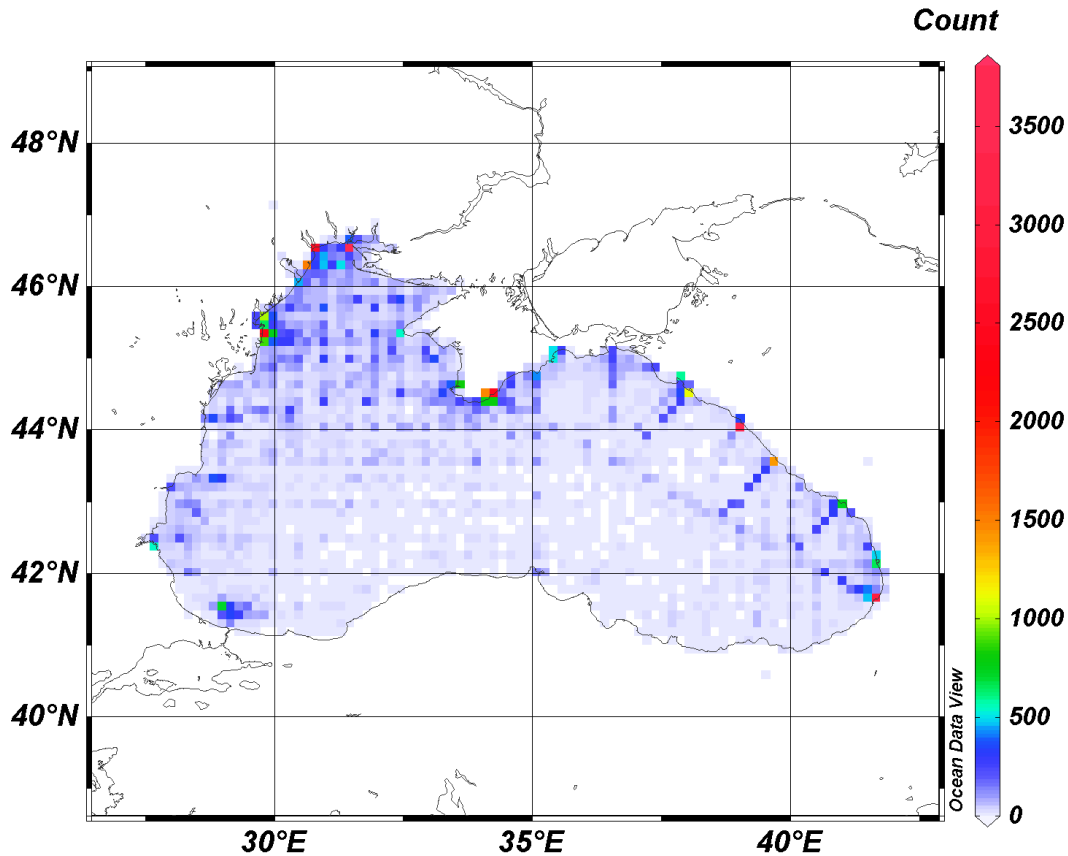
Data distribution by Years



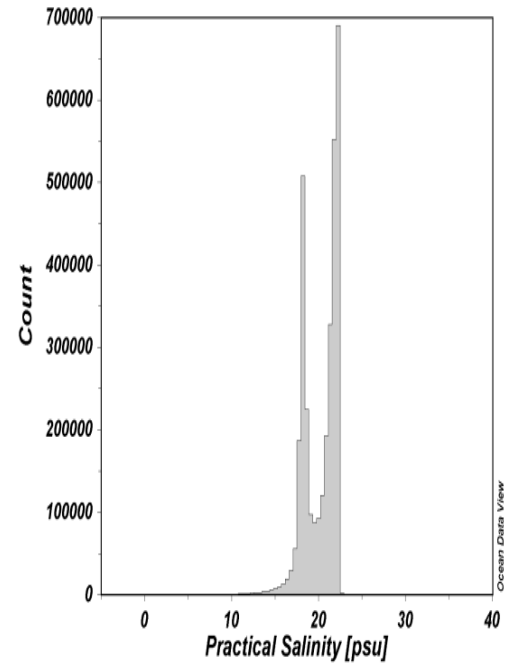
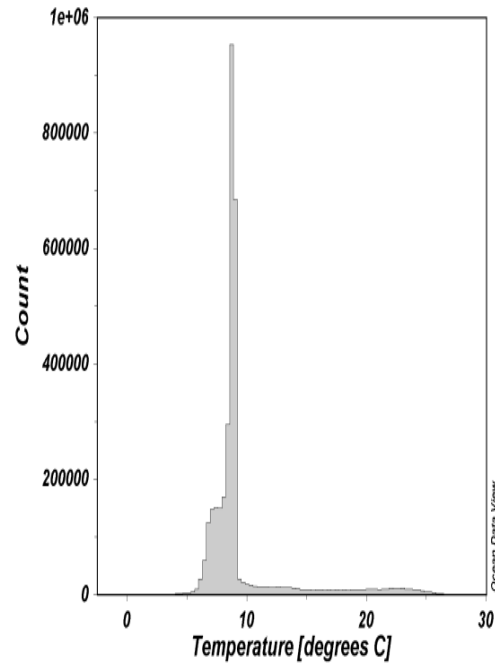
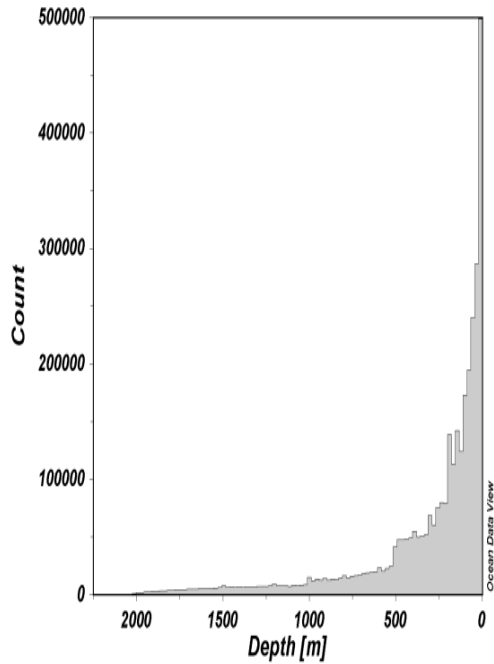
Data distribution by season



Spatial data distribution



Data distribution histograms



Data quality flagging statistics



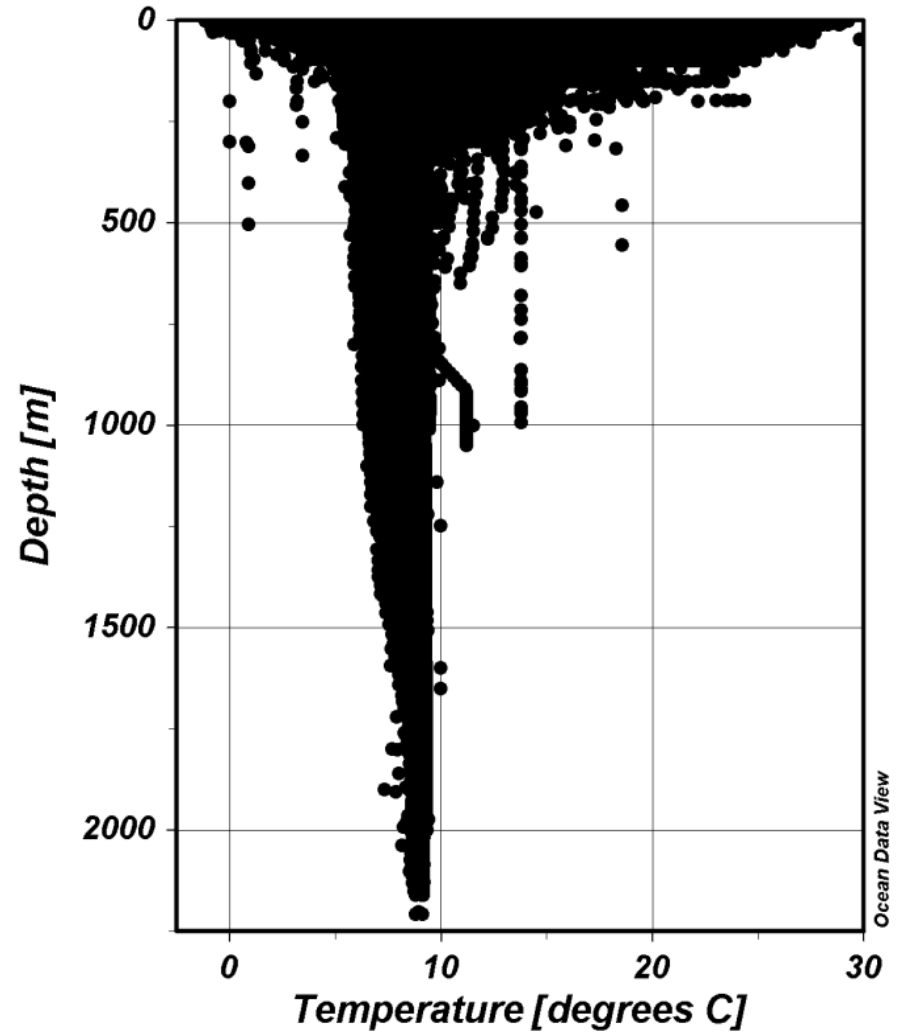
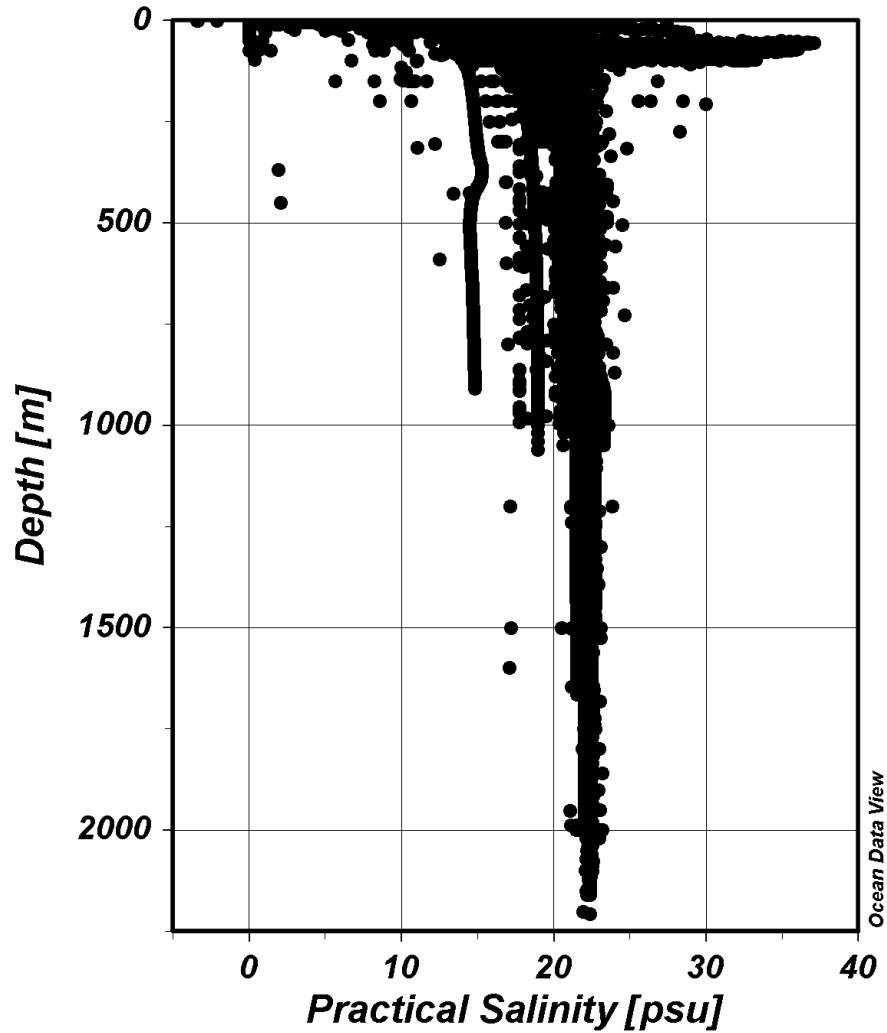
QF	0	1	2	3	4	Total
	not checked	good	probably good	probably bad	bad	
Depth	479258	2885854	1	13175	81898	3460186
	13.85%	83.40%	0.00%	0.38%	2.37%	100.00%
Temperature	526632	2868689	45934	19649	55	3460959
	15.22%	82.89%	1.33%	0.57%	0.00%	100.00%
Salinity	515381	2839665		18959	802	3374807
	15.27%	84.14%		0.56%	0.02%	100.00%

Challenges (1)

- Not all existing QF can be trusted
 - Simple range check revealed that definitely wrong values (e.g. negative Salinity values or missing value indicators such as “9999” or “-99”) are flagged as Good (!)

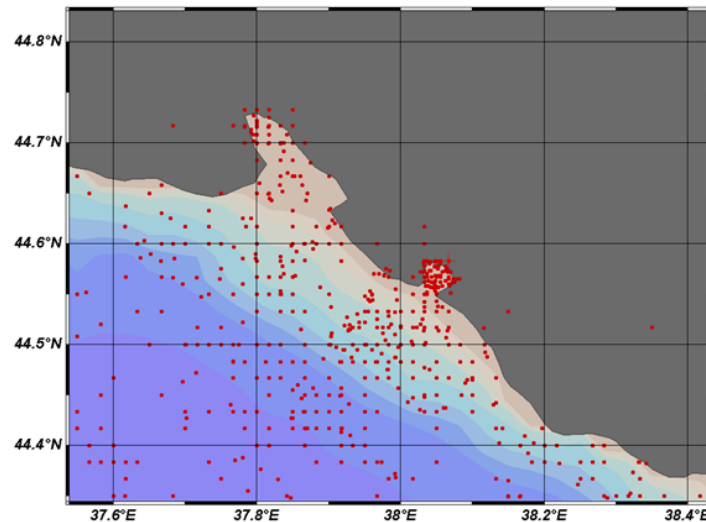
Depth		Temperature		Salinity	
Min	Max	Min	Max	Min	Max
0	2212	-1.5	30	0	37
46		64		1120	

Scatter plots



Challenges (2)

- Stations on land - total 2065



- Stations with T-S data without depth: 967
- Stations where profile depth is greater Bot.Depth: 1794
- Duplicates: 5700 (from them 467 are absolute duplicates)

Future QC phase

- Complete checking and excluding duplicates
- Q: shall we exclude from the Black Sea T-S dataset the stations in river estuaries and firths (limans)? (there a lot in NW)
- Perform QC of non-flagged data
- Perform checking of flagged data
 - Identify datasets having not correct QC flagging
 - Perform new QC for these dataset
- Send consolidated instructions to data providers on improving their datasets