SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

MyOcean QC feedback analysis

Christine Coatanoan

SeaDataNet Plenary Meeting, Lucca, 26-27 September 2013

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

Aggregated dataset – different steps

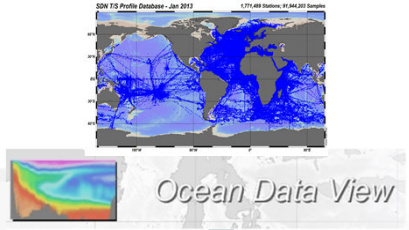
CDI extraction of all the datasets SeaDataNet Licence

(CDI : Discovery and data access)

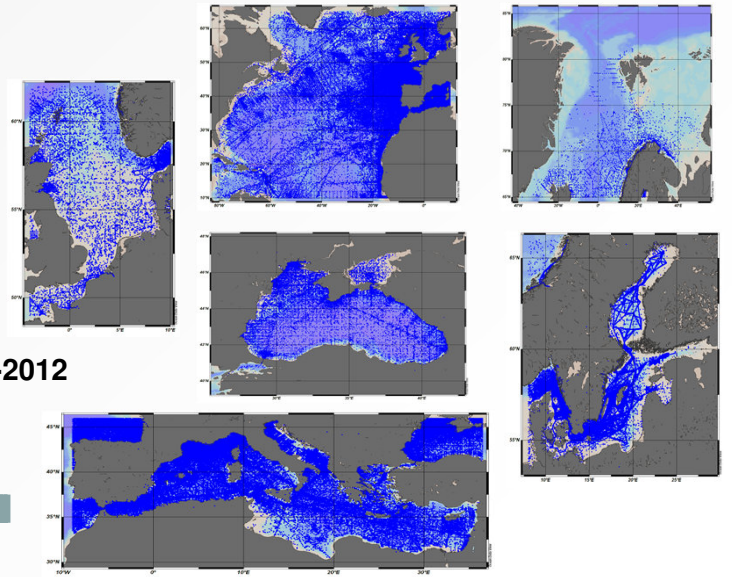
AWI

Check of duplicates

Correction by EDMO source



Regional subsets collection



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)**

1990 to 2011

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

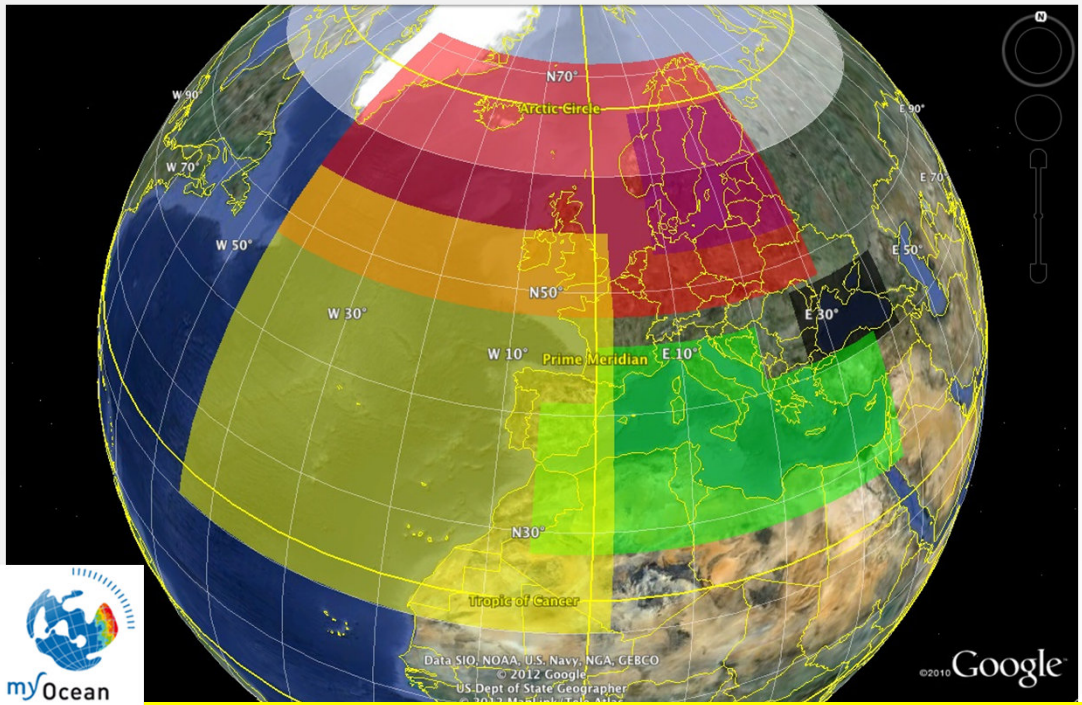
PAN-EUROPEAN INFRASTRUCTURE
FOR OCEAN & MARINE DATA
MANAGEMENT

SeaDataNet

myOcean

Aggregated dataset – North Atlantic

	West	East	North	South
Arctic	-180	180	90	60
South West Shelf	-50	9	60	19
Baltic	31	8	66	53
Black Sea	42	26	47.5	40
Mediteranea2	20	0	45.6	41
Mediteranea1	37	-5.5	41	28
North West Shelf	31.5	-45	71.5	48
Global	-180	180	90	-90



Baltic Sea : N°53-N°66, E°008-E°031
North Sea :N°49-N°62, W°004 – E°010
Arctic:N°65-N°80, E°010-E°075
Med Sea : N°30-N°46, W°009-E°036.5
Black Sea: N°40-N°47, E°026.5-E°042
Atlantic: N°10-N°65, W°082-E°010

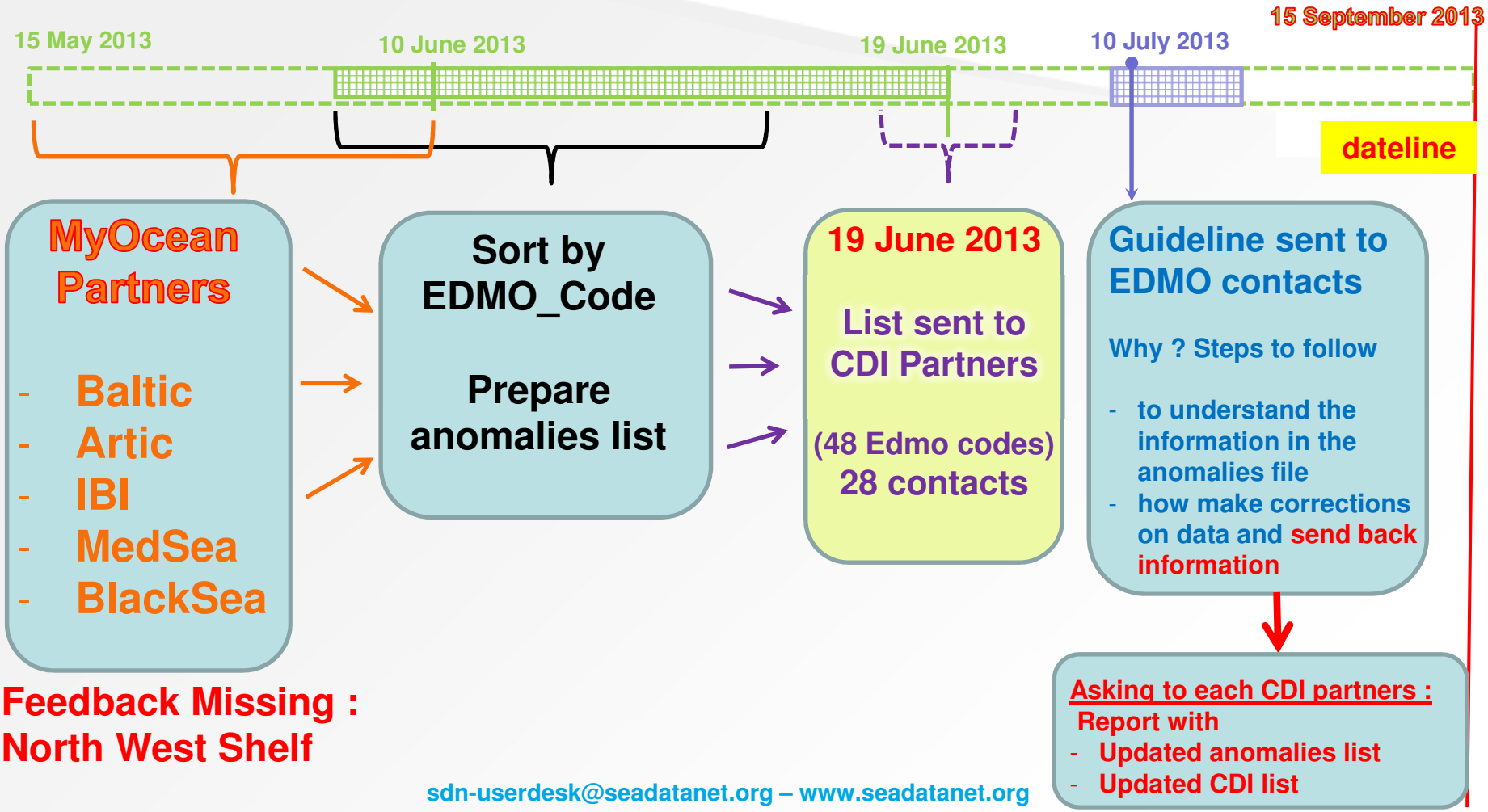
SeaDataNet

MyOcean

myOcean

February/March 2013

Exchanges between partners : schedule

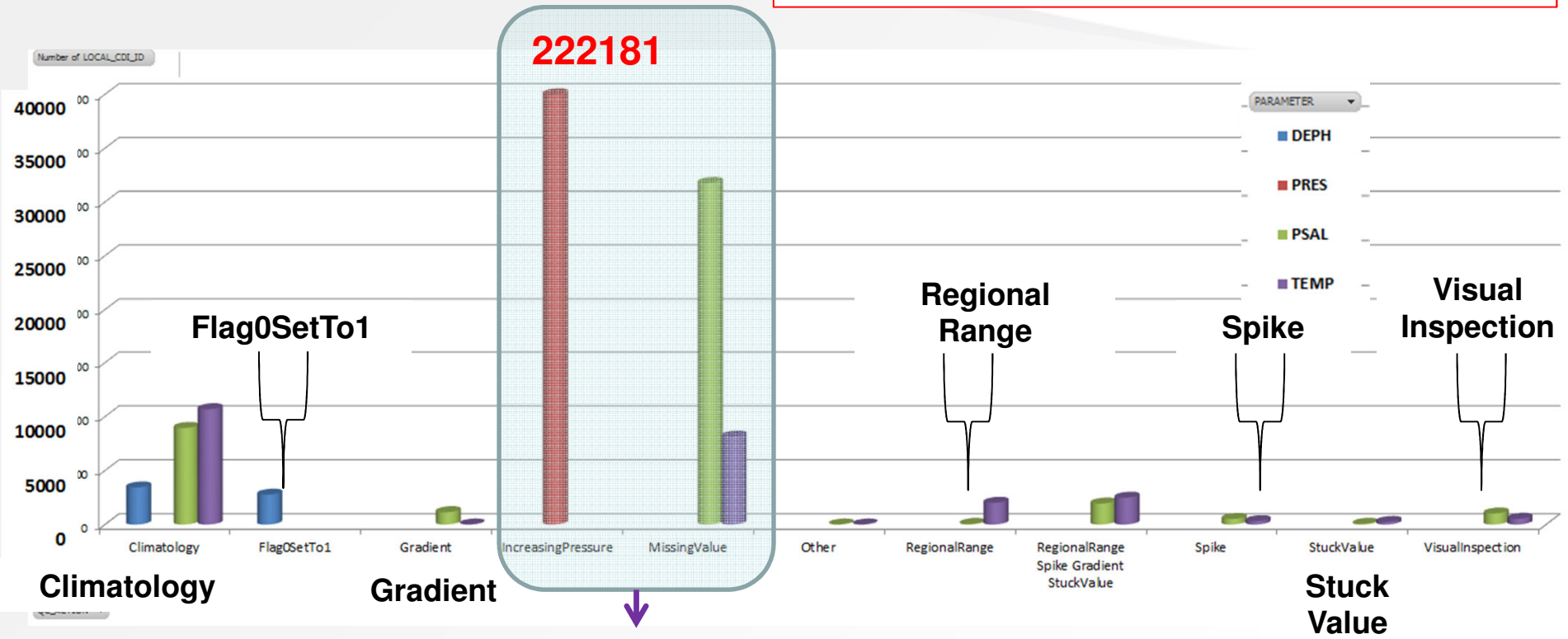


Feedback Missing :
North West Shelf

SeaDataNet

Climatology : the data is out of the climatology envelope
Gradient : a gradient anomaly
Increasing pressure : a non increasing depth
Missing Value : a missing value with a flag not equal to 9
Regional Range : a value out of the regional range
Spike: spike
StuckValue : a constant profile
Visual Inspection : MyOcean visually finds an anomaly

Types of anomalies



Increasing pressure & Missing Value
Not real anomalies ? may not be due to a problem with the data (see after)

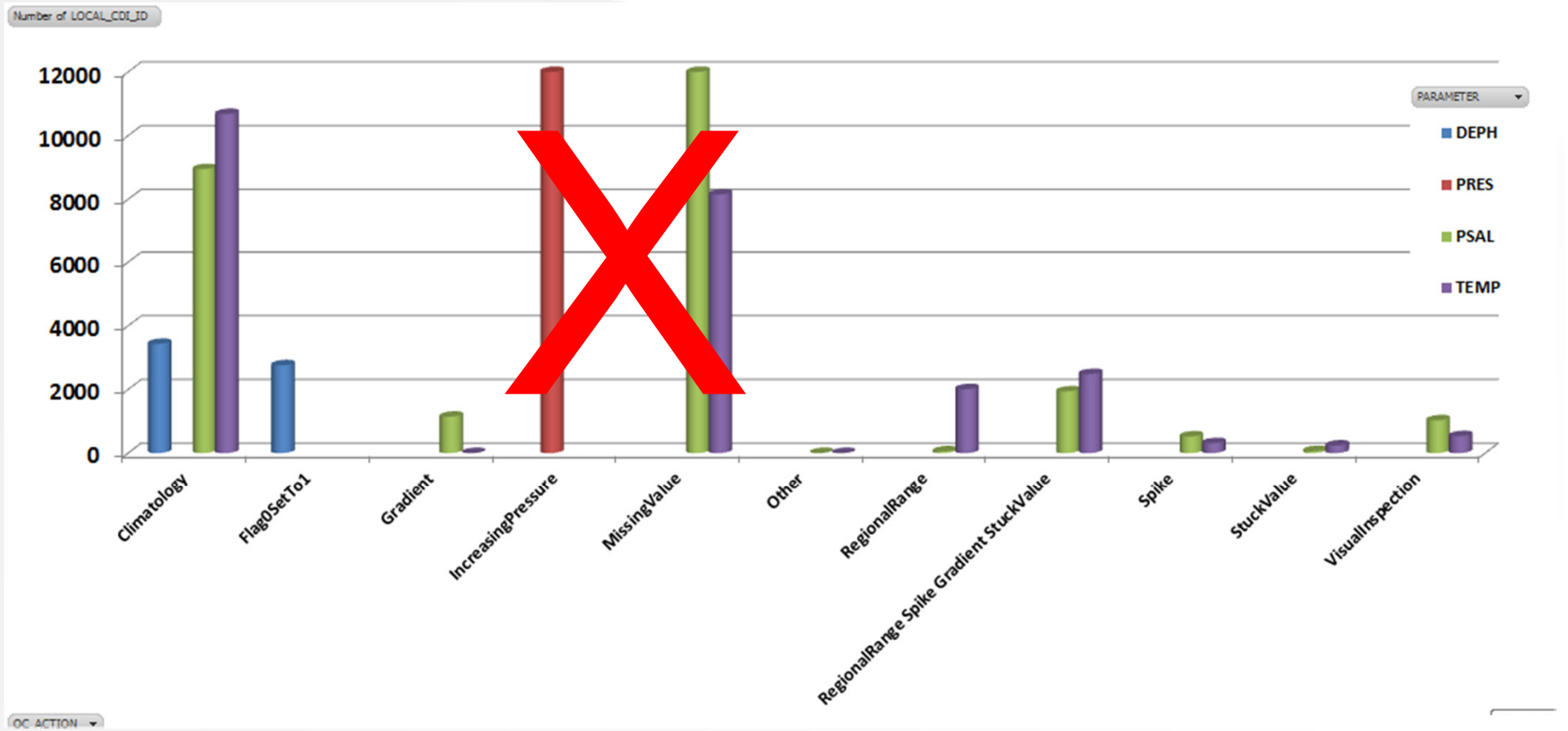
Increasing pressure & Missing Value

- **QC action "Missing Value"**

QC action "Missing Value" 0 to 9 came from a bug in ODV during export, this detected errors must not be taken into consideration

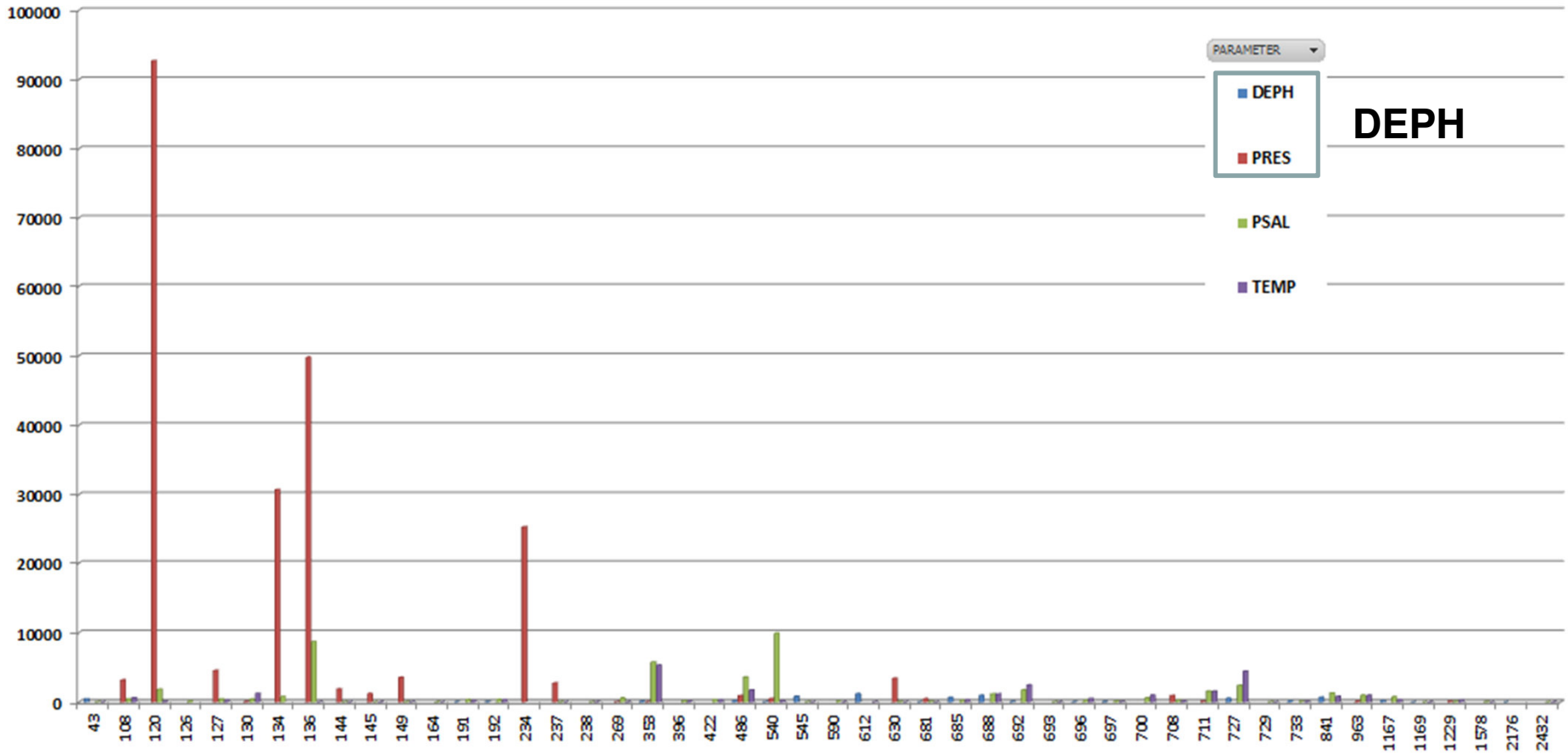
- **QC action "Pressure increasing" can have several origins:**

- If the type of platform is Bottle : it can be normal to have several samplings at the same depth, so in that case there is no anomaly
- It can also be a problem of duplicates CDIs (in some cases when data providers forwarded to Maris multiple updates of CDIs in a short time frame with overlaps), and in that case also it is not a true anomaly
- It can be a true error of non-increasing depth/pressure and then value should be flag to 4


Number of LOCAL_CDI_ID
12000
10000
8000
6000
4000
2000
0
PARAMETER
DEPH
PRES
PSAL
TEMP
Climatology
FlagOSetTo1
Gradient
IncreasingPressure
MissingValue
Other
RegionalRange
RegionalRange Spike Gradient StuckValue
Spike
StuckValue
VisualInspection
QC ACTION

SeaDataNet

NUMBER OF ANOMALIES BY EDMO_CODE AND BY PARAMETERS



DEPH

EDMO_CODE (48)
=> 28 contacts

What is expected from each CDI partner ?

- The anomalies list updated with NODC comments following this table
- The list of updated CDI :

LOCAL_CDI_ID	EDMO_CODE	PLATFORM_CODE=CRUISE
FI35199101301_00050_H10	486	PRIMO-0 21/03
FI35199443005_25900_H10	486	MBP-FRONT 1994
FI35199502002_00870_H10	486	EUROMARGE
FI35199706005_0K010_H10	486	PELMED 97
FI35199845001_00260_H10	486	BIODYPAR 1

- A report with some informations

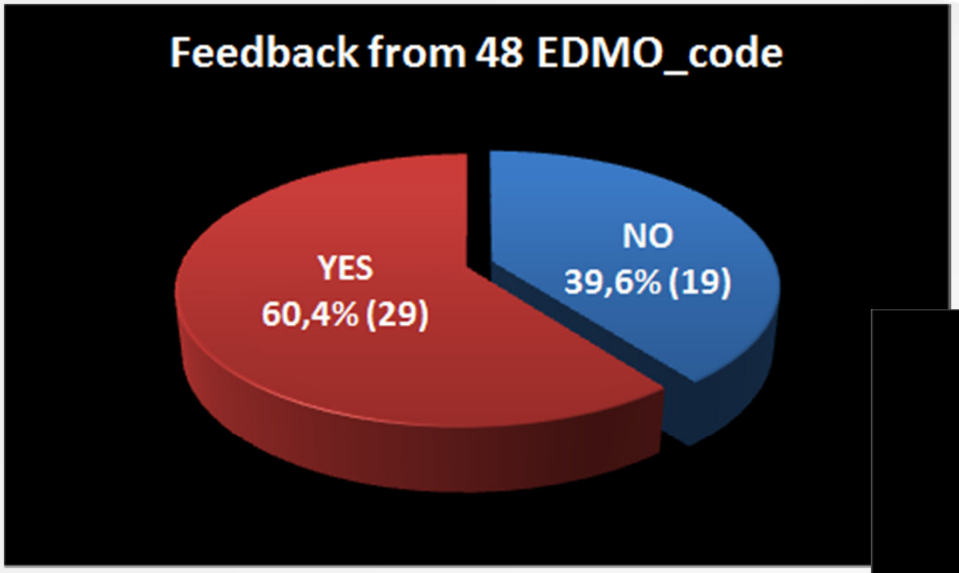
- List of errors and number

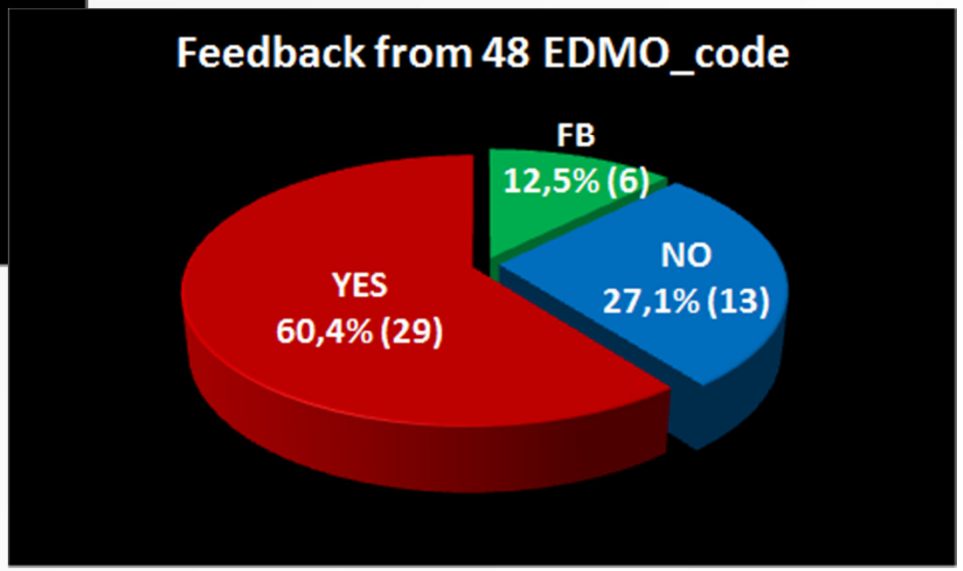
QC_Action	Number of anomalies detected by MyOcean	Number of true anomalies	%
Climatology	20	0	0.0
Gradient	328	1	0.3
IncreasingPressure	544	25	4.6
RegionalRange	42	0	0.0
Spike	148	42	18.3
StuckValue	28	0	0.0
VisualInspection	1	1	100.0
Total	**1111**	**69**	**6.2**

- Details on why corrections have not been taken into consideration, etc……

Column	Description	Comment
LOCAL_CDI_ID	*cdi_identifier*	Partner local CDI identifier, Information from CDI
EDMO_CODE	EDMO_CODE of the organization distributing the data	Information from CDI
PLATFORM_CODE=CRUISE	CDI cruise_name	
STATION_DATE_START	Date at which the station starts	
STATION_DATE_STOP	Date at which the station ends	
UPDATE_DATE	Date of the control done by MyOcean partners	
PARAMETER	PARAMETER exported from ODV (TEMP, PSAL, DEPH or DEPTH [sometimes PRES when MyOcean partners have changed name])	
QC_ACTION	As described in Introduction, to define the type of anomalies (spike, gradient, missing value, etc)	
OLD_QC	QC from original dataset	
NEW_QC	QC suggested by MyOcean (see Annex I)	
VERTICAL_REFERENCE_START	Level at which starts the anomaly in the profile	
VERTICAL_REFERENCE_STOP	Level at which stops the anomaly in the profile	
AGREE WITH THE SUGGESTED CORRECTION (YES/NO)	Fill with Yes/No if you agree/disagree with the corrections suggested by MyOcean	
NODC COMMENT	Column to be added to your file in order to put some information about your our opinion about the suggested correction (agreement, disagreement, explanation if necessary)	
DETAILS	Column to be added to your file in order to put more information about suggested correction	

Some statistics (based on 19th September)





⇒29 Feedbacks on 48 EDMO contacts

Feedback from NODC without result files

In progress but we need corrections and updated CDI for the next release (beginning of october) !!

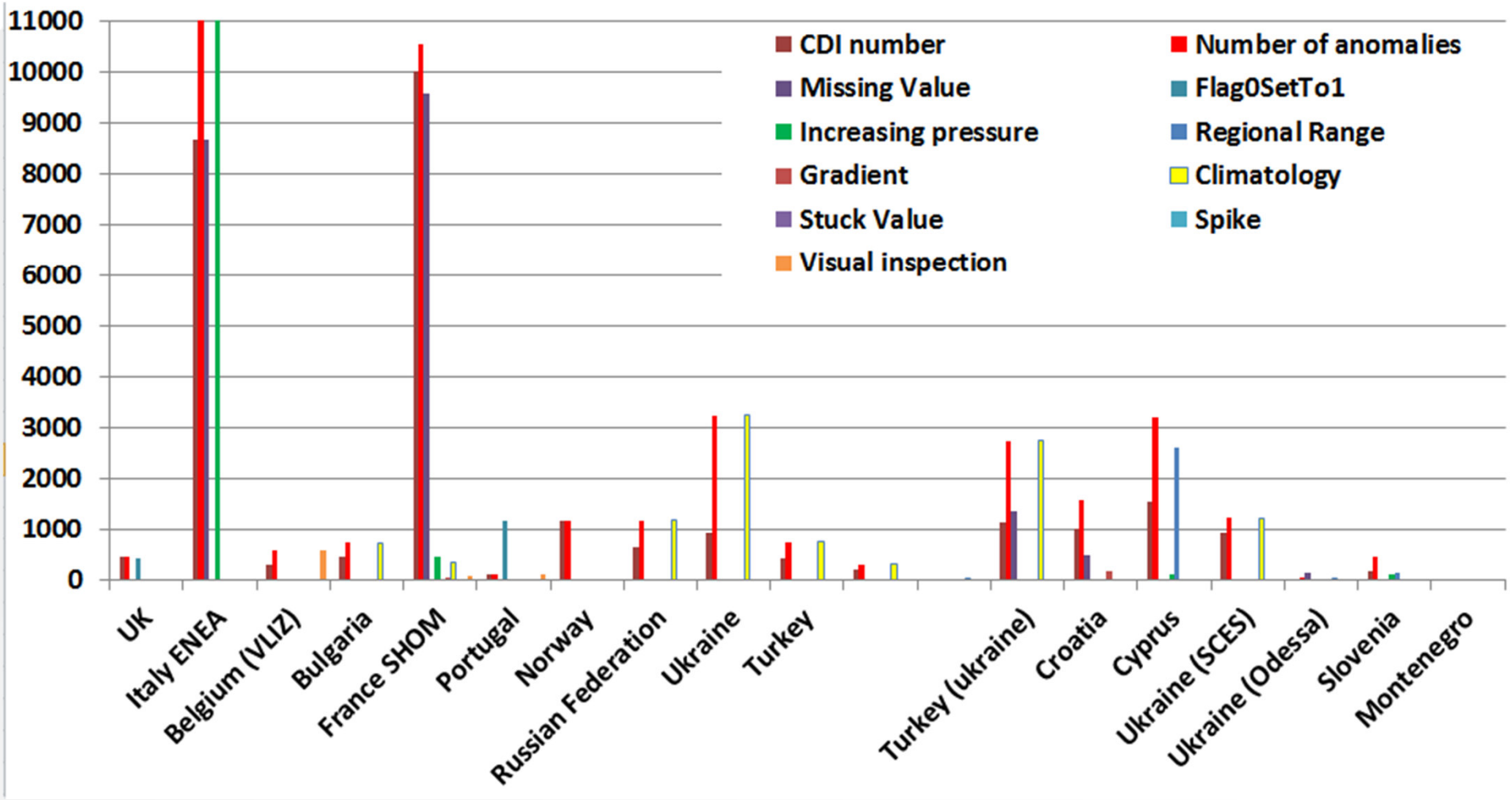
Contact	EDMO_CODE	CDI number	Number of anomalies	Regional Range	Gradient	Climatology	Stuck Value	Spike	Visual inspection
Italy Giuseppe	136	8680	58447	0	0	0	2	0	1
Bulgaria	192	456	737	0	0	0	0	0	0
France SHOM	540	10021	10567	10	54	349	0	2	95
Portugal	590	112	125	0	0	0	0	0	125
Cyprus	711	1546	3199	2609	0	0	0	0	0
Slovenia	1229	188	444	142	22	0	12	12	0

Number of anomalies (included : missing value & increasing pressure)

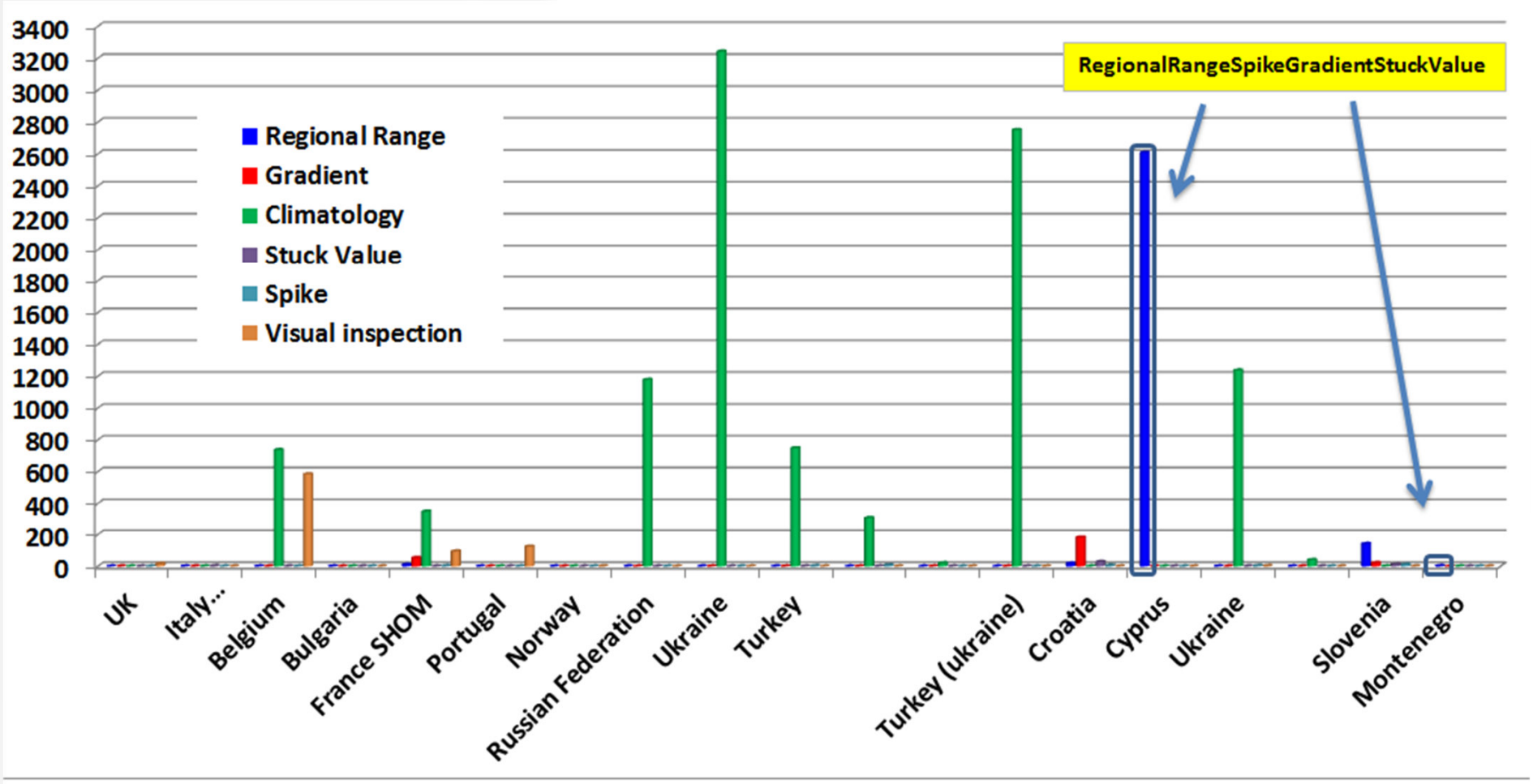
Feedback from NODC : any answer

Contact	EDMO_CODE	CDI number	Number of anomalies	Regional Range	Gradient	Climatology	Stuck Value	Spike	Visual inspection
UK	43	455	458	0	0	0	0	0	17
Belgium	422	315	585	0	0	737	0	0	585
Norway	612	1173	1174	0	0	0	0	0	1
Russian Federation	685	634	1178	0	0	1178	0	0	0
Ukraine	688	929	3243	0	0	3243	0	0	0
Turkey	696	433	750	0	0	748	0	2	0
	733	222	316	0	0	309	0	7	0
	2176	21	21	0	0	21	0	0	0
Turkey (ukraine)	841	1159	2751	0	0	2751	0	0	0
Croatia	700	1019	1590	19	182	0	29	5	0
Ukraine	1167	929	1242	0	0	1237	0	3	2
	1169	15	41	0	0	41	0	0	0
Montenegro	2432	1	2	2	0	0	0	0	0

Number of anomalies (included : missing value & increasing pressure)


11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
CDI number
Missing Value
Increasing pressure
Gradient
Stuck Value
Visual inspection
Number of anomalies
Flag0SetTo1
Regional Range
Climatology
Spike
UK
Italy ENEA
Belgium (VLIZ)
Bulgaria
France SHOM
Portugal
Norway
Russian Federation
Ukraine
Turkey
Turkey (ukraine)
Croatia
Cyprus
Ukraine (SCES)
Ukraine (Odessa)
Slovenia
Montenegro

Anomalie types *(removing missing value & increasing pressure)*



EDMO_CODE	Name	Country	Collating centre	SDN2 contact
43	British Oceanographic Data Centre (BODC)	UK	British Oceanographic Data Centre (BODC)	Lesley Rickards <ljr@bodc.ac.uk>
422	Flanders Marine Institute (VLIZ)	Belgium	Management Unit of the North Sea Mathematical Models (RBINS-MUMM)	klaas.deneudt@vliz.be tjess@vliz.be
612	Institute of Marine Research - Norwegian Marine Data Centre (NMD)	Norway	Institute of Marine Research - Norwegian Marine Data Centre (NMD)	helge.sagen@imr.no
685	P.P.Shirshov Institute of Oceanology, RAS	Russian Federation	P.P.Shirshov Institute of Oceanology, RAS	shiganov@ocean.ru
688	Southern Scientific Research Institute of Marine Fisheries and Oceanography	Ukraine	Marine Hydrophysical Institute	khaliulin.alexey@nodc.org.ua
696	Institute of Marine Sciences, Middle East Technical University	Turkey	Institute of Marine Sciences, Middle East Technical University	devrim@ims.metu.edu.tr dtezcan@gmail.com
700	Institute of Oceanography and Fisheries	Croatia	Institute of Oceanography and Fisheries	dadic@izor.hr ivankovic@izor.hr
733	Sinop University, Fisheries Faculty	Turkey	Institute of Marine Sciences, Middle East Technical University	devrim@ims.metu.edu.tr dtezcan@gmail.com
841	Marine branch of Ukrainian Hydrometeorological Institute	Ukraine	Marine Hydrophysical Institute	devrim@ims.metu.edu.tr dtezcan@gmail.com
1167	Ukrainian scientific center of Ecology of Sea (UkrSCES)	Ukraine	Marine Hydrophysical Institute	khaliulin.alexey@nodc.org.ua
1169	Odessa National I.I.Mechnikov University	Ukraine	Marine Hydrophysical Institute	khaliulin.alexey@nodc.org.ua
2176	Ankara University	Turkey	Institute of Marine Sciences, Middle East Technical University	devrim@ims.metu.edu.tr dtezcan@gmail.com
2432	Institute of Marine Biology (IMBK)	Montenegro	Institute of Marine Biology (IMBK)	biokotor@gmail.com klodiz2005@yahoo.com

Feedback from NODC (files provided)

Corrections or no correction ?

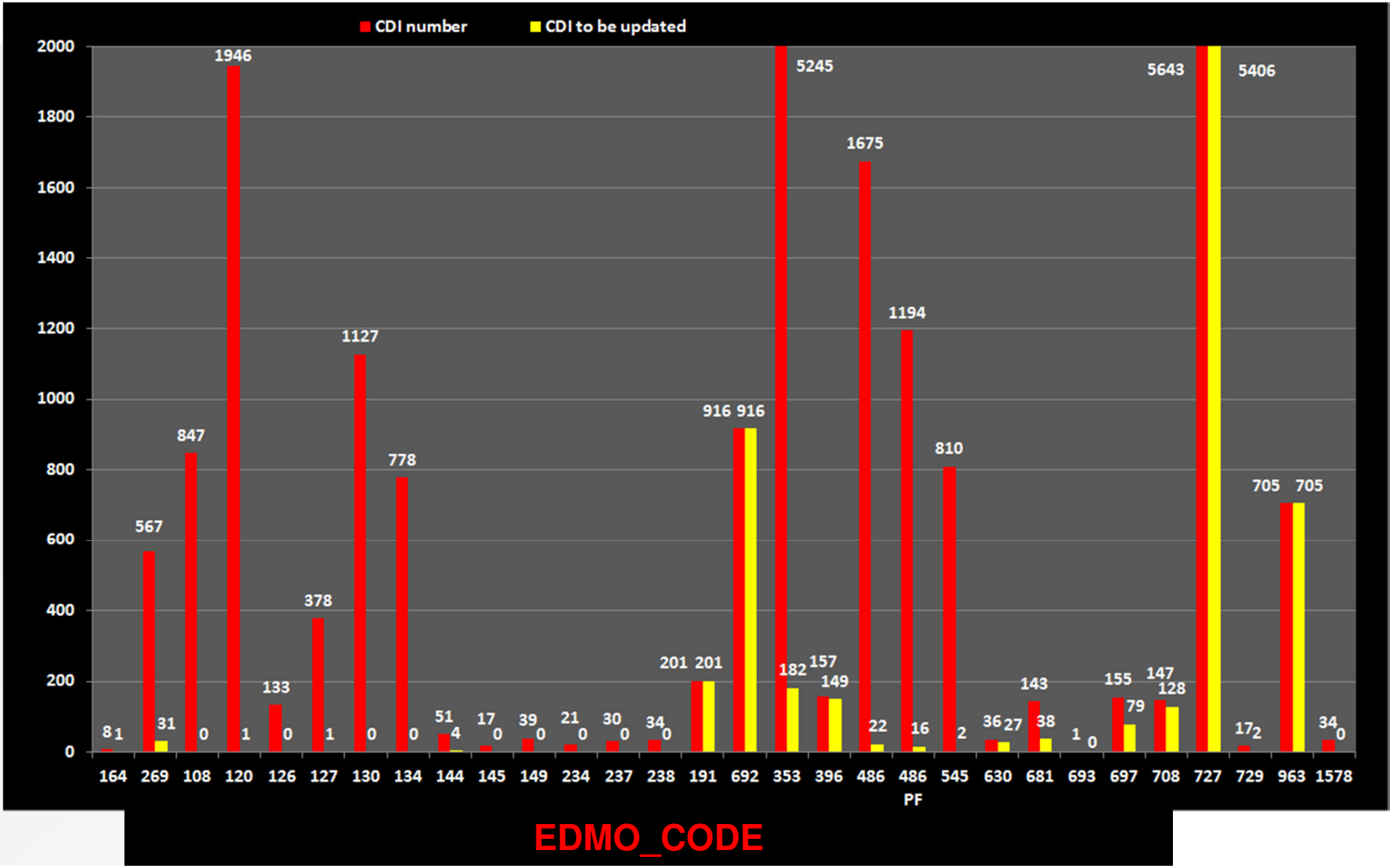
- already done in CDI
- no correction needed
- correction applied

SeaDataNet

PAN-EUROPEAN INFRASTRUCTURE FOR OCEAN & MARINE DATA MANAGEMENT

Work done on 29 answers

23055 CDI – 7911 to be updated



In terms of corrections (on 29 answers)

- 2111260 anomalies (all actions including 19669 for missing value & 171707 for increasing pressure)
- **18739** anomalies without missing value & increasing pressure

- 14405 corrections done (on **18739**) => 76,87%

Most are from Climatology/Visual Inspection/Spike

In term of corrections (on 29 answers)

In term of correction number : Major anomalies

Climatology (11929 corrections), Gradient (1516), Visual Inspection (437), Spike (317), Stuck value (189), Regional range (17)

In term of % by type of anomalie:

- Climatology 95,97%
- Gradient 63,22%
- Visual Inspection 57,27%
- Stuck Value 42,28%
- Spike 2,55%
- Regional range 0,90%

Anomalies which are not anomalies

- Increasing pressure : when the type of platform is bottle, we can have duplicates (or more) levels

- Spike : again, for bottle platform, we can have

- one measurement detected as spike
- or 2 measurements (ex. 1 : surface, 2 : at 100m), then mostly surface measurement is detected as a spike

Gradient – no real anomalie

stations even in the River

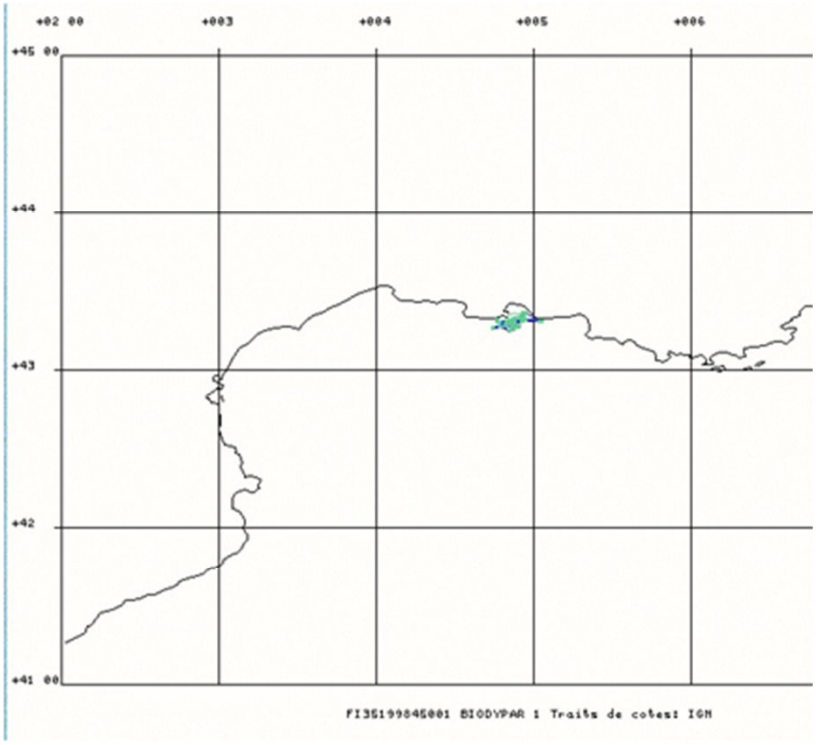


Figure 1 – Station located in the Rhone plume, and corresponding valid Salinity profiles

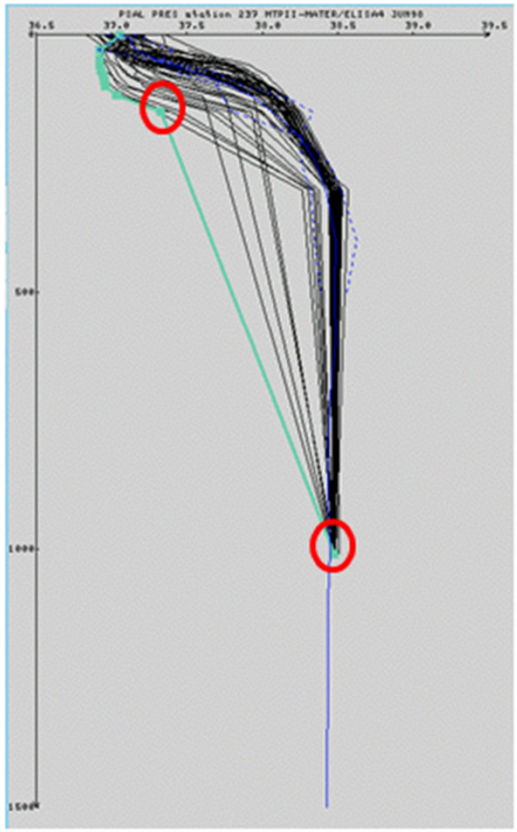
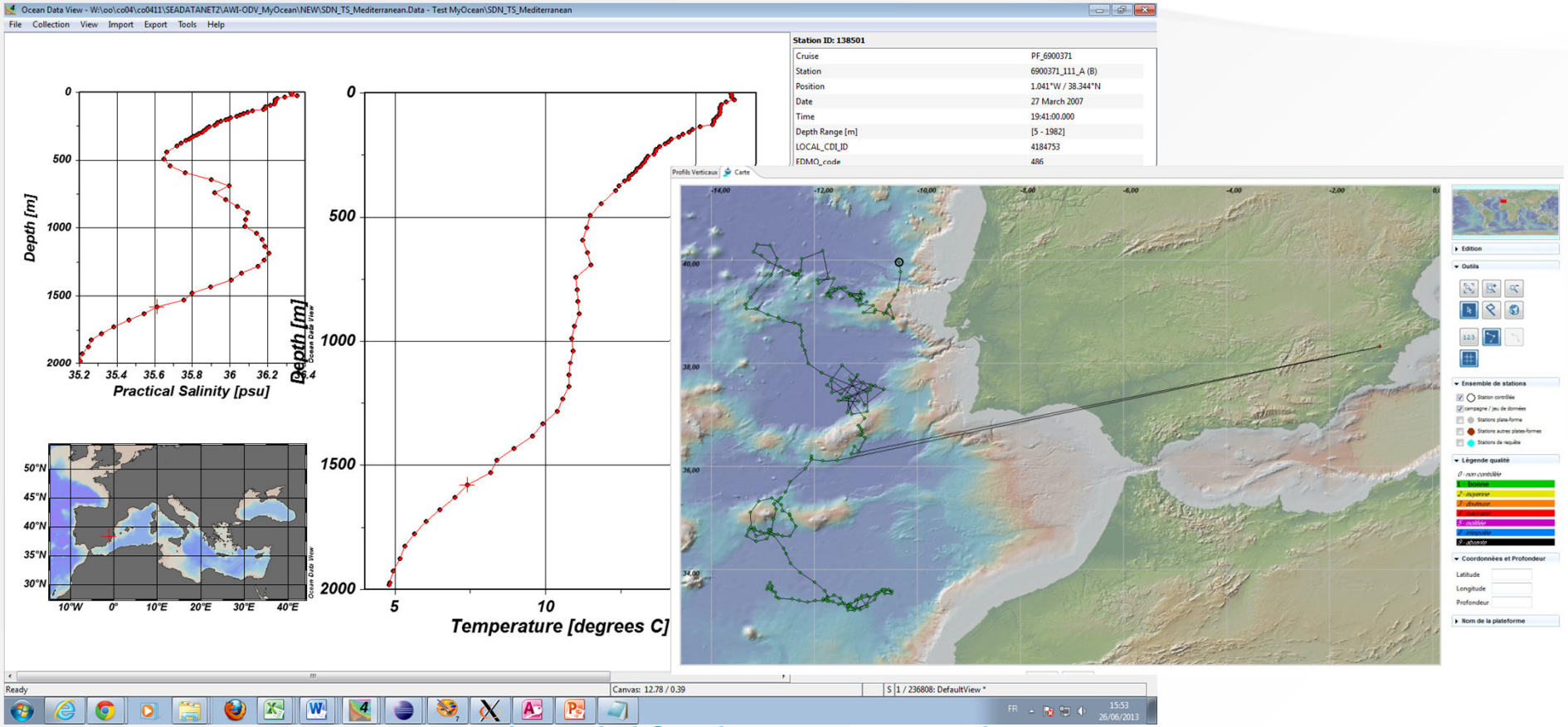


Figure 2 - Gradient detected by MyOcean between the 2 red circles

Real anomalies

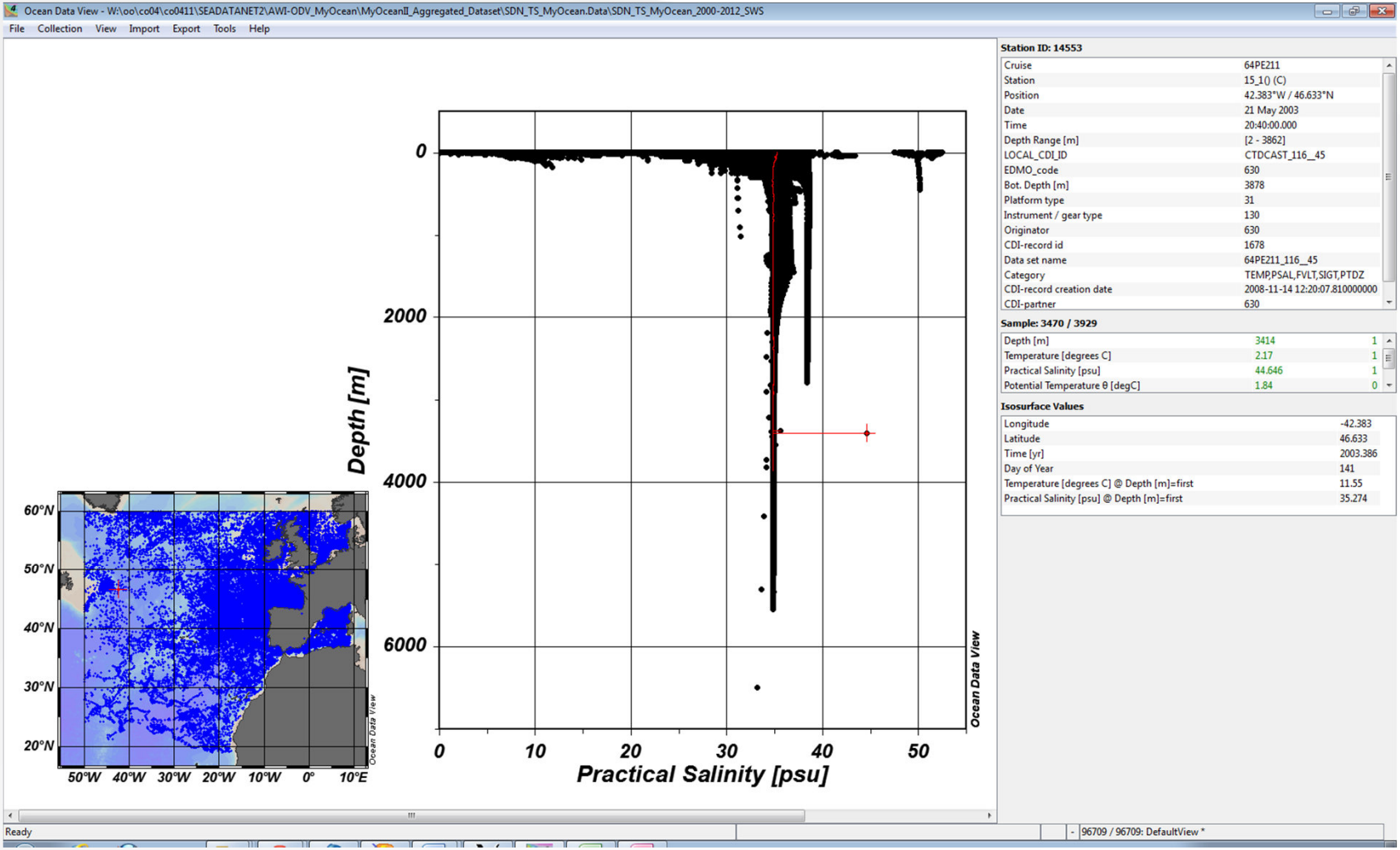
Ex. TEMP Regional range

4184753	486	PF_6900371	2007-03-27T19:41:00Z	2007-03-27T19:41:00Z	2013-05-28T22:26:13Z	TEMP	RegionalRange	1	4	1334,64	1982,12



Real anomalies

Ex. PSAL Spike – Regional range



Feedback TO MyOcean

- Real anomalies (spike, bad values, regional range…) – corrections status
- Missing value QC 0 due to error in ODV export
- Increasing pressure : duplicated levels [due to a duplicate of profiles in the .lst – CDI export twice of the same local_cdi_id] or not real duplicated (bottle platform)

Feedback TO MyOcean about MyOcean QC procedure

- Take care of the type of platform (when available), specially for bottle to not identify again increasing pressure or spike
- Take care of the coastal stations : automatic tests seem not to be enough successful in those oceanic areas