



## 4th EuroGOOS DATAMEQ meeting

The 4<sup>th</sup> EuroGOOS DATAMEQ Working group meeting was held at the EuroGOOS Office in Brussels/Belgium from the 10<sup>th</sup> October at 2pm to the 11<sup>th</sup> September 2013 at 1pm.

The objectives of the meeting was to work on Term of Reference 1,2,3,4 and 5 (see annexe é)based on the development taking advantage of the developments performed within MyOcean , SeaDataNet and EMODnet-Physics in close link with the EuroGOOS ROOSes

1. *Revise what has been set up since 2010 and identify the improvements that are needed especially concerning the links between Real-Time and Historical Data streams ( MyOcean-SeaDataNet- EMODnet-Physics)*
2. *Reach an agreement on RTQC procedures developed between MyOcean and SeaDataNet for bio-geochemical data (MyOcean, JERICO feedback)*
3. *Define Action plan for next 2-5 years*

Sylvie Pouliquen opened the meeting with the summary of the recommendation that were elaborated in 2010 and endorsed by the EuroGOOS Assembly . These recommendations are summarized in the meeting report that can be found in the DATAMEQ section of the EuroGOOS WWW sites ([http://www.eurogoos.org/content/documents.asp?menu=0050000\\_000015\\_000000](http://www.eurogoos.org/content/documents.asp?menu=0050000_000015_000000)). She showed the new membership of the group that evolved with time as project and expertise needed were changing. The new membership is in Annex1 as well as the list of participants. 15 participants from 5 among the 6 ROOSes attended the meeting.

## **1 What are the progresses through projects related to data access that can benefit EuroGOOS**

### ***1.1 Data exchange and integration within MyOcean (S Pouliquen)***

Sylvie Pouliquen presented the developments made in MyOcean and MyOcean2 jointly with the EuroGOOS ROOSes to develop a Global portal (Coriolis at Ifremer/France) and 6 regionalportals ( Arctic at IMR/Norway, Baltic at SMHI/Sweden, North West Shelves at BSH/Germany, South West Shelves at Puertos del Estado/Spain, Mediterranean at HCMR/Greece and Black Sea at IOBAS/Bulgaria) that aggregate both near real time quality controlled observations ( within a few hours from acquisition), and aggregated historical products ( 1990-2012) in partnership with SeaDataNet2, for forecasting and reanalysis activities.

She pointed out the importance of the near real time QC performed on the data provided by the ROOS partners, and of the validation done at basin scale to assess the consistency of the data prior to assimilation in the models, as a significant number of anomalies are detected through these processes. This is one reason why web services on data available remotely is not presently enough to ensure the quality of the aggregated products requested by modellers.

While setting up these portals, MyOcean In situ TAC partners faced a certain number of difficulties in particular while building historical data set and integrating data coming from SeaDataNet, ROOS

partners and the NRT data streams. She suggested the following recommendations to be discussed with the group:

- Use the same standards at list at European level based on SeaDataNet standards whenever exists : unique Platform code, Provider code
- Provide a list of mandatory metadata that would be exchanged with the data
- Define a common strategy to handle duplicates and get the “best copy” out of the network. Provide error-bar and accuracy of the measurement with the data
- Improve QC on data by efficient reporting provider when an anomaly is detected
- Provide clear guidelines to institutes who want to exchange the observation they manage to connect to either SeaDataNet network or MyOcean/EuroGOOS portals or both

These points were further discussed during the work plan elaboration discussions.

## 2 Linking Historical and Real Time station within EMODNET-Physics

Peter Thijsse from Maris and Antonio Novellino from ETT presented the progress made within the EMODnet-Physics project to provide visibility to Real-time and Historical fixed buoys platforms and Ferrybox data provided on the portals developed jointly by MyOcean INSTAC partners and the ROOSes on one side, and SeaDataNet network of NODCs on the other side.



They showed what was working well and highlighted the difficulties and weaknesses of the present situation and identified the following challenges:

- Challenge to ensure that NRT data flow is in sync with data as available at original providers, concerning available parameters per station and quality of time series in time
- Challenge to ensure that metadata in NRT index catalogue is consistent in time to support linking the harvested data from the FTP exchange to the other services (Monthly NRT files, EDIOS catalogue and SeaDataNet CDI service) by means of station name / id
- Challenge to expand the number of exchanged stations to all stations as operated by the existing data providers
- Challenge to ensure that SeaDataNet will get a more complete coverage of all available long term series and that these are populated in the CDI service
- Challenge to ensure that used station identification metadata are consistent in time and in sync with the ID metadata as in use in the FTP exchanges to support linking the SeaDataNet CDI entries to the other services (daily NRT exchange, monthly NRT index, and EDIOS service)

In the frame of JERICO and EMODnet\_Physics it was observed that there at present 900 fixed monitoring stations are identified in Europe of which so far about 500 are included in the EuroGOOS near real-time exchange. Moreover shortcomings are observed in the exchange of existing stations considering completeness of parameters. To improve the present situation EMODNet-Physics is preparing an inventory of the stations integrated presently in the system identifying the discrepancies between the provider holdings and the data available both on the ROOS portals and in CDI-SeaDataNet. Then the EMODnet-Physics partners would like to work with ROOS on how to solve the issues and improve the overall system.

Noting that EMODnet Physics has a budget of 500.000 Euro for EuroGOOS members for undertaking such efforts, following interactions with EMODnet Physics were presented :

- EMODnet Physics will deliver a document with analysis per data provider (Oct 2013):
  - Possible shortcomings present NRT exchange
  - Possible extra stations for inclusion in NRT exchange
  - Overview of SeaDataNet archives
- The group recommends further action by data providers and distributors to improve the situation considerably by:
  - EuroGOOS members coordinated by ROOSes undertake follow-up (Oct 2013 – June 2014):
    - Check and correct shortcomings NRT exchange
    - Check and arrange adding extra stations NRT exchange
  - EuroGOOS members together with SeaDataNet colleagues undertake (Oct 2013 – Oct 2014):
    - Populating the SeaDataNet CDI service with full coverage of long term validated data sets for monitoring stations (historic and present)
    - Thereby paying attention to using consistent station names / ids
  - Regular communication between ROOSes and EMODnet-Physics partners about progress and to safeguard high quality

It was pointed out that a handbook that would explain to providers how to exchange their data holdings with MyOcean/ROOS portals and SeaDataNet would be useful, that sharing common user information directory between MyOcean & SeaDataNet would streamline the downloading facilities managed at EMODnet level.

Enhancing visibility of original observation providers in these portals was seen as an important topic to involve more partners. It was suggested by Gisbert Brietbach to include in the metadata information linked to observation files links to the available web service that allow to discover, view and download using WFS, SOS. Finally Thomas Loubrieu presented Oceanotron, that will be deployed in all ins-situ TAC portals that will provide viewing and download services from the portals and plan to be also used by SeaDataNet.

### **3 Towards recommendations for BIO-Geochemical RTQC procedures**

Within MyOcean and JERICO FP7 projects, activities started in 2011 to define Real Time quality control procedure for some biogeochemical parameters measured and transmitted to shore in near real time. The first priority has been put on Chla and O<sub>2</sub> parameters but the goal is to extend such procedures to nutrient, PCO<sub>2</sub> and PH when there will be NRT data streams.

Presently the V2.0 version of the procedure is mature enough to be circulated within the DataMEQ working group and the target is to arrive to endorsement of a first document at next EuroGOOS meeting. Kai Sorensen from NIVA provided scientific background to highlight the variability of such measurements and also the importance of the calibration made in laboratory prior to deployment as well as the importance of water samples to correct the data in delayed mode ( This is possible on vessels but not on floats or gliders that are autonomous platforms).

**Action:** *It was pointed out that in JERICO , some good deliverables on quality assurance procedures have been written for Glider, Ferrybox and Mooring and It was suggested to the EuroGOOS Technical WG to study the possibility to turn such deliverable into an EuroGOOS recommendation document.*

Kai then presented the proposed tests:

- global range test,

- regional tests : the threshold are in some areas difficult to define due to the lack of historical data
- spike and gradient : it's a 3 step methods to avoid removing natural variability
- Bio-fouling test that should flag data as suspicious when it fails

An Instrument and parameter relation test was envisaged but it is difficult to implement in Near Real time

To move forward towards the goal to have an NRTQC document for Chlorophyll and Oxygen ready for the EuroGOOS meeting the following actions were decided:

- NIVA to circulate the present document to DATAMEQ-WG , MyOcean In Situ TAC partners
- Sylvie to send Argo NRTQC procedure for Oxygen and the 1st Bio Argo meeting report to NIVA
- Suggestion for NIVA to attend bio-argo meeting 14-15 October in Liverpool to link to what is done in Argo .

## 4 Define Action plan for coming 2-5 years as a set of recommendations

The meeting continued with a brainstorming session to elaborate the recommendation the group wanted to be presented at next EuroGOOS meeting in November 2013 to be endorsed as action plan for EuroGOOS through projects and ROOSes activities for the coming years

The group reinforced the importance of an open data policy for EuroGOOS, for a data exchange that relies as much as possible on standards when they exist. They recognized the fact that this system should be developed in collaboration with European Data Exchange projects but that these services needed to be endorsed by the ROOSes and extended to fit their own purposes.

In particular they identified the following major projects to interact with:

- MyOcean for near real-time data stream
- SeaDataNet and MyOcean for building historical products for reanalysis purposes
- SeaDataNet for standards improvements, scientific validation, and long term stewardship and access of data archives
- JERICO for coastal network expertise
- EmodNet-Physics for interoperability between SeaDataNet and MyOcean as well as the other EmodNet portals.

The group identified 3 main categories of actors

- Data provider, together with the platform operator, ensures that the data are acquired correctly and that enough metadata are attached to the observations to trace the processing history.
- Data distributor collects data from data providers and provides homogeneous distribution services to users.
- Regional centre aggregates data from data distributors or providers and provides an integrated service to users. For some parameters, it checks the data consistency over the region and flags the suspicious data. It can build value added products for users.

The group recommends to continue extending the regional portals that are set up through projects (SeaDataNet , MyOcean, JERICO, PERSEUS...) to the parameters and platforms necessary for the ROOSes activities.

To improve the quality of the data exchanged in near real time:

- The group recommends updating the DataMEQ RTQC procedures for T&S, Current and Sea-Level especially for coastal data.
- The group recommends publication of RTQC procedure for O<sub>2</sub>/Chl recommendations based on the work developed within MyOcean, JERICO, GROOM and EuroArgo.
- The group recommends further coordination between platform operators for setting up new technology in order to meet operational oceanography needs (e.g. HF Radar).

To enhance the interoperability between real-time and delayed mode data streams

- The group recommends harmonization of some essential metadata at EuroGOOS level:
  - Platforms should be referenced using a unique Platform code. The group recommends use of the existing Platform vocabulary (C17/C174) as provided by the SeaDataNet infrastructure and managed by the ICES/NOAA/Seavox platform governance group that assigns unique platform identifiers, including all necessary attribute information. SeaDataNet will work with EuroGOOS, the ROOSes and

JCOMMOPS to increase the content of the platform vocabulary, particularly for those platform classes that are currently not well represented.

- Providers should be referenced using a unique Provider code. The group recommends use of the EDMO vocabulary managed by SeaDataNet, extended to include monitoring institutions in addition to research institutes.
- A list of mandatory metadata should be exchanged with the data information, including: how data have been processed, provenance, link to web services where available...
- No data should be distributed without X&Y&Z&T reference.
- The group recommends definition of a common strategy to handle duplicates and get the “best or master copy” out of the network.
- The group recommends provision of quality information about the measurement (error-bars, accuracy, standard deviation)

To improve interaction and involvement of new providers, the group recommends enhancement of provider and user interaction, including:

- Organization of periodic regional meetings to facilitate provider and user interaction
- Provision of clear guidelines to institutes through a handbook for institutes that want to exchange the observation they manage
- A feedback loop set up by regional centres between provider and distributor during regional assessment
- Provide material to provider that they could use at local level, improve their visibility from regional centre,
- Continuation of work on data citations (DOI etc)
- Setting up an efficient update process between regional centers and providers
  - To ensure that NRT data flow corresponds to data as available at original providers
  - To ensure that SeaDataNet will have more complete coverage of all available long term series

To enhance the quality of the acquired data the group recommends that the Technical working group publishes a best practice handbook including recommendations for inclusion of the necessary metadata from platform to data centres.

Finally for historical data exchange, the group maintained the 2010 recommendations

- The group recognized the importance of an up-to-date catalogue of the existing observing systems in Europe such as EDIOS, and recommend a collaboration with the ROOSes to update such a catalogue in a distributed and automated manner
- Historical in situ datasets are important for reanalysis activities carried out within EUROGOOS. Therefore the group recommends SeaDataNET to set up periodic data sets providing the “master copy” (removing duplicates) for a core set of parameters (starting T&S).
- As a counterpart SeaDataNet asked the ROOSes to provide feedback on the quality of the yearly update to help SDN improving these products that will be delivered officially on a four year basis fully qualified

The meeting ended at 13h00

## 5 Annexe1 DATAMEQ members

*Updated the 1<sup>st</sup> August 2013*

*Participants in Black, excused in orange*

**Chair** Sylvie Pouliquen

**EUROGOOS** Patrick Gorringe

### **BOOS and NOOS representatives**

BSH Kai Soetje

SMHI Thomas Hammarklint

UKMET **Matthew Martin, Simon Good**

CEFAS Kate Collingridge /**Rodney Forster/Jon Rees**

### **IBI-Roos representatives**

Puertos Del Estado Marta de Alfonso

AZTI Julien Mader

### **MOON representatives**

Enea Giuseppe Manzella

HCMR: Leonidas Perivoliotis

### **Arctic representative**

IMR **Helge Sagen , Sjur Ringheim Lid, Henning Wehde**

### **Black Sea representative**

IOBAS **Veselka Marinova**

### **Other program representatives**

Argo/Gosud/OceanSITES/GROOM **Thierry Carval**

SeaDataNet **Dick Schaap/Peter Thijsse**

EMODnet-Physics Antonio Novellino

Interoperability Tools Thomas Loubrieu

Ferrybox **Wilhelm Petersen/Gisbert Breitbach**

ICES **Neil Holdsworth/Hjalte Parner**

RTQC-BIO (MyOcean JERICO) Kai Sorensen / NIVA

## **Terms of Reference**

*Validated at EuroGOOS annual 2011*

1. Develop an overall concept for the management of EuroGOOS observation data taking into consideration data management systems which are developing within GMES and JCOMM
2. Identify, in consultation with the EuroGOOS Task Teams and OOS/N, as appropriate, the type of observations which can be made available either in real-time or in delayed mode
3. Propose the most effective ways to make observation data readily available for operational purposes in a sustained matter
4. Propose mechanisms to ease access to delayed mode observation data in cooperation with NODCs, keeping aware of the progress in SeaDataNet
5. Draft a minimum set of standards for data quality control which is related to observation data collection, processing and exchange procedures

Each TT or OOS/N should appoint 2 persons to represent them in the working group and promote internal coordination. The DATA-MEQ chair will liaise with the JCOMM Data Management Program Area coordinator.