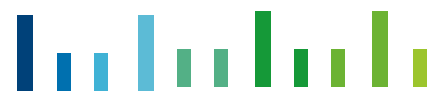


Joint European Research Infrastructure network for Coastal Observatories



D# 5.8 (Near) Real-time Data Management Handbook V 2

Grant Agreement n° 262584

Project Acronym: JERICO

Project Title: Towards a Joint European Research Infrastructure network for Coastal Observatories

Coordination: P. Farcy, IFREMER,

jerico@ifremer.fr, www.jerico-fp7.eu:

Authors: R. Nair, L. Petit de la Villeon

Involved Institutions: OGS and IFREMER

Version and Date: V2 – 10-03-2015





TABLE OF CONTENTS

1. DOCUMENT DESCRIPTION	5
2. EXECUTIVE SUMMARY.....	5
3. INTRODUCTION.....	7
4. MAIN REPORT	9
4.1. The current JERICO data management structure and principal data flows	9
4.2. Near real-time data.....	11
4.2.1 Description of the MyOcean near real-time data management infrastructure	11
4.2.2 Uploading the main JERICO near real-time data types	13
4.2.2.1 Fixed platforms.....	13
4.2.2.2 Ferryboxes	13
4.2.2.3 Gliders	14
4.2.2.4 Fishing vessels (Fishery Observing Systems).....	17
4.2.3 Useful contacts for near real-time data submission.....	18
4.3 Downloading the main JERICO near real-time data types	20
5. CONCLUSIONS.....	21
6. LIST OF ACRONYMS.....	22



1. Document description

REFERENCES

Annex 1 to the Contract Description of Work (DoW) version of the 22 Feb. 2011

Document information	
Document Name	Data Management Handbook, Version 2, D#5.8
Document ID	
Revision	
Revision Date	
Author	Rajesh Nair, Loic Petit De La Villeon
Security	

History			
Revision	Date	Modification	Author

Diffusion list				
Consortium beneficiaries	X			
Third parties				
Associated Partners				
other				

This document contains information, which is proprietary to the JERICO consortium. Neither this document nor the information contained herein shall be used, duplicated or communicated by any means to any third party, in whole or in parts, except with prior written consent of the JERICO Coordinator.

The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.





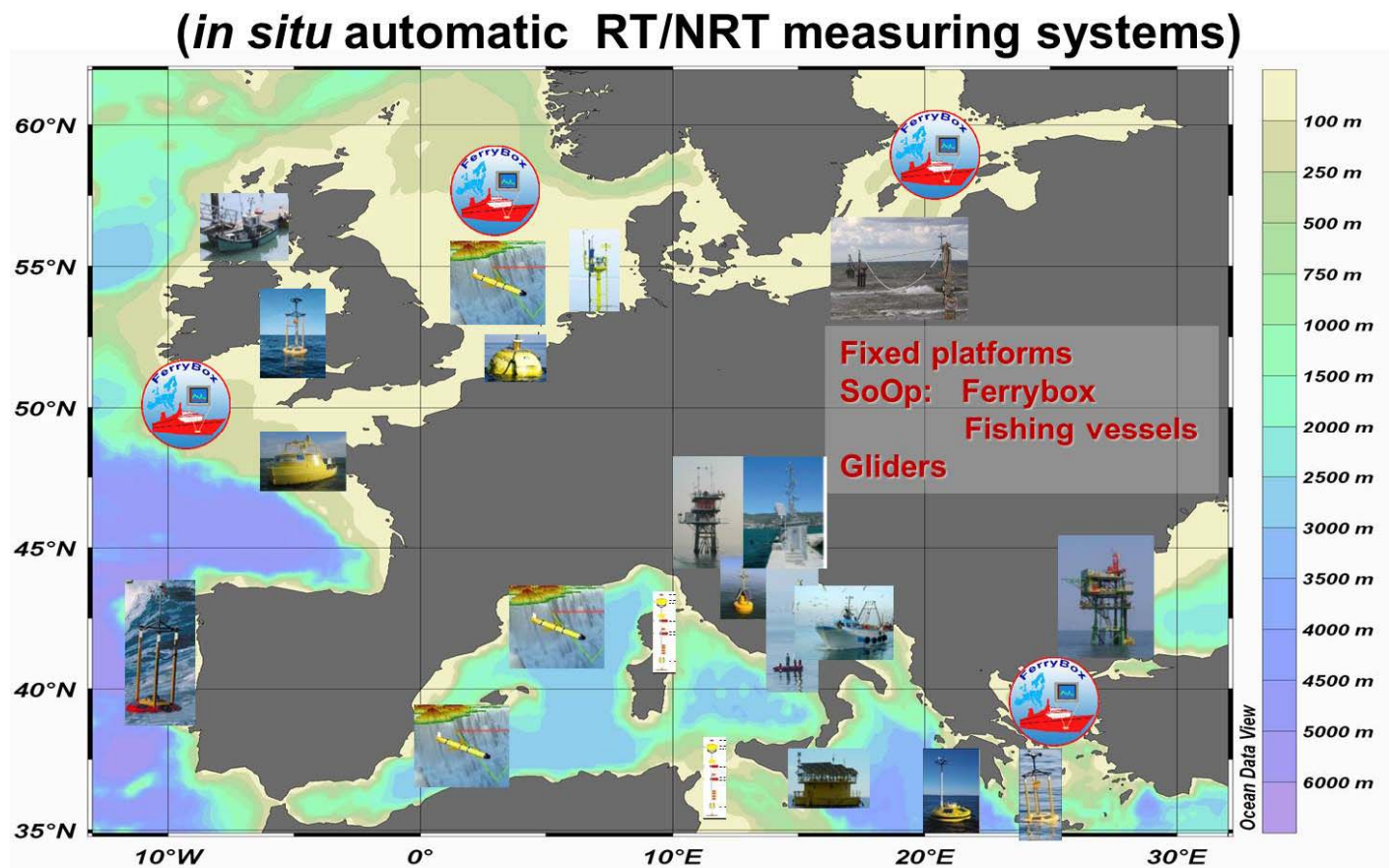
2. Executive summary

The success of the coastal observatory network created through the JERICO project depends on the reliability, accessibility and easy distribution of data coming from the participating observing systems. This necessarily implies the use of a standard platform and common procedures for handling the data produced by the network. The present document constitutes the second version of the JERICO Near Real-time Data Management Handbook. It describes the general JERICO data management structure and policy, and provides partners with practical advice and useful information on how to manage their real-time data within the framework of the project. The document also contains references and links to the basic and most important online documents needed for implementing the established procedures. The guidelines presented here are those that have been followed by partners for handling their real-time data during the JERICO Service Data Access period that began in January 2013.



3. Introduction

The FP7 JERICO project embodies the first operational network of coastal observatories - assemblages of distributed sensor systems with extensive coverage - on a truly pan-European scale. The network is strikingly heterogeneous, marked by a surprising richness in the range of parameters handled, the frequency and spatial distribution of measurements, equipment maintenance practices, and the quality assurance schemes employed for sensors and data. All the coastal observatories forming the JERICO network share the same goal: to help detect and investigate coastal processes in a timely fashion, and provide crucial operational data for planning, assessment, mitigation, and model assimilation and validation purposes. The current composition of the JERICO network of coastal observatories is summarized in Fig 1.



Source: Sparnocchia et al. (Journal of Operational Oceanography, submitted)

Fig 1: Current composition of the JERICO network of coastal observatories showing the main kinds of real-time (RT) or near real-time (NRT) observing platforms in use.

As shown, the network mainly comprises four kinds of observing platforms: fixed or station-keeping structures, Ferryboxes on ferries, fishing vessels equipped with sensors mounted on fishing gear, and gliders; in more than a few cases, operators manage complex systems incorporating more than one of these platforms. The measuring instruments and sensors on the platforms can differ widely in number, kind, scope, and technical configuration from platform to platform, and from operator to operator. The sets of targeted parameters also vary considerably across the network, though salinity, temperature and pressure (depth) are nearly always measured. Table 1 provides a breakdown of the principal parameters being handled in the JERICO project.

Table 1. The JERICO parameter list; only the core parameters are presented.

	PARAMETER	CORE	OPTIONAL
PHYSICAL	Salinity	!!	
	Temperature	!!	
	Turbidity	!!	
	Sea level	!!	
	Surface waves	!!	
	Surface currents		!!
	Chlorophyll-a	!!	
	Turbidity	!!	
	CDOM		!!
	Noise Passive Acoustic Listener (PAL)		!!
CHEMICAL	Dissolved nutrients		!!
	Dissolved oxygen	!!	
	CO2 partial pressure	!!	
	Contaminants		!!
	pH or Alkalinity		!!

Source: Sparnocchia et al. (Journal of Operational Oceanography, submitted)

The JERICO network is geared towards ensuring continuous, timely access to coastal observations. The emphasis is on assuring valid data streams for coastal data at the transnational level in near real-time to ensure prompt availability and in delayed-mode for more deferred use.

4. Main Report

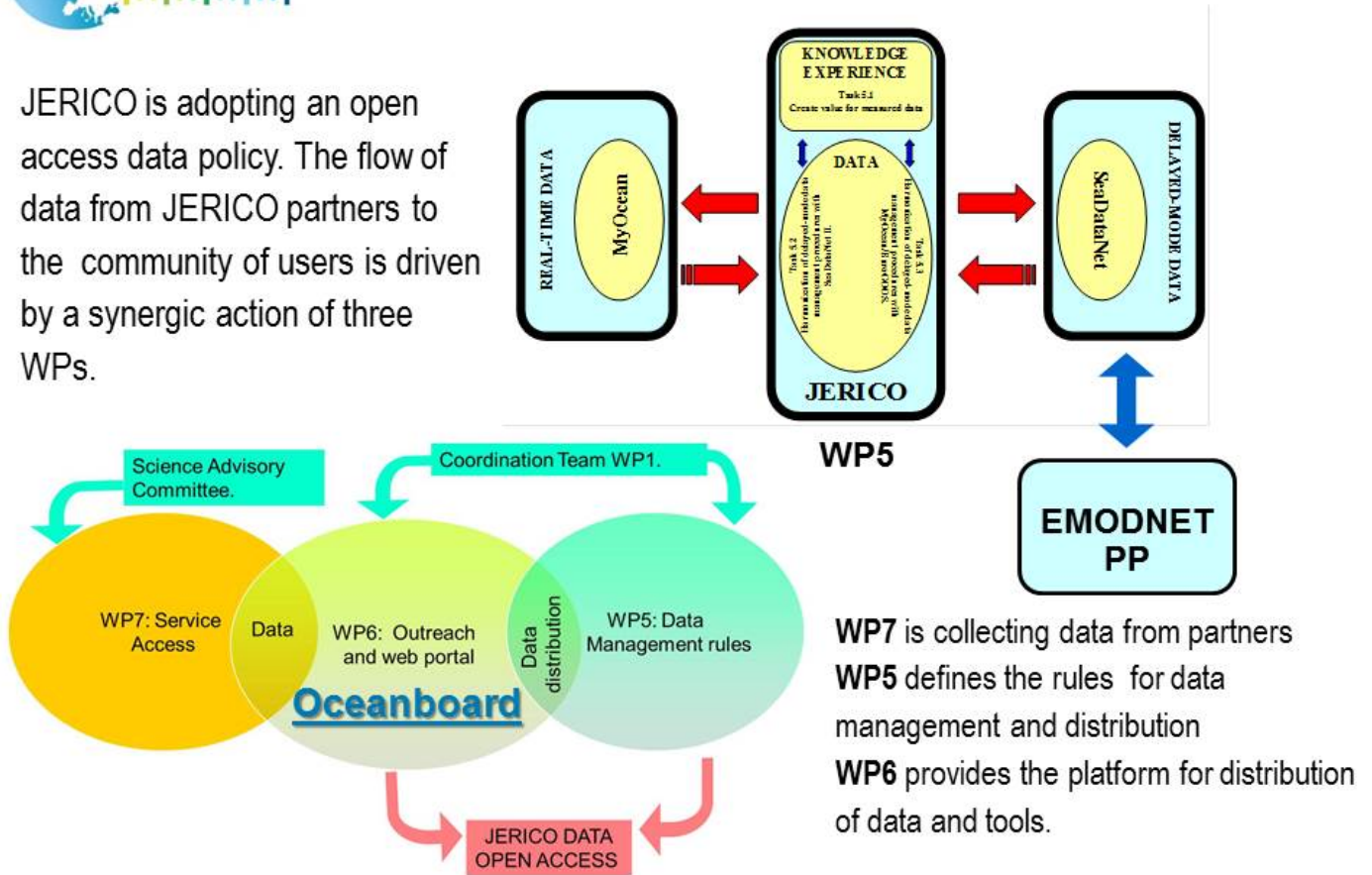
4.1. The current JERICO data management structure and principal data flows

The general approach to data handling that has been adopted by the JERICO project is outlined in Figure 2. It is based on an “open” data access policy, following the conventional European outlook in these matters. The flow of data from partners to the community of users concerns three Work Packages: WP7 is collecting data from partners, WP5 defines the rules for data management and distribution, and WP6 provides the platform for the distribution of data and related tools.



Access to JERICO data

JERICO is adopting an open access data policy. The flow of data from JERICO partners to the community of users is driven by a synergic action of three WPs.



Source: Sparnocchia et al. (submitted)

Fig 2: Outline of the JERICO data management scheme, featuring the driving actions and main information flows.



The data management structure is built on the principle of “using what exists”. Suitable partnerships have been created with ongoing European data management initiatives so as to minimize possible duplication of efforts. Thus, there has been no dedicated development of a specific data management structure for JERICO. Instead, the use of, and integration with, already available data management infrastructure has been pursued, a strategy consistent with the policy behind SeaDataNet and MyOcean, the major ongoing European initiatives for the establishment and coordination of infrastructures for the management and distribution of marine data and products.

The JERICO data management framework for delayed-mode data uses the SeaDataNet infrastructure while the near real-time data are being routed through MyOcean. There is continuous interaction with SeaDataNet, MyOcean, EuroGOOS and EMODnet to enable the seamless integration of the JERICO data stream into these two established European marine data management infrastructures. Many difficulties remain to be overcome, though, especially in relation to parameters and data types that are not actually being handled in SeaDataNet and/or in MyOcean.

Some specific dataset indexing procedures were developed to help ensure compatibility with MyOcean Thematic Assembly Center requirements and assist non-NODC data providers in using the SeaDataNet infrastructure. An appropriate indexing and data distribution scheme was created. A distinct tag, an explicit JERICO index, was designed to easily recognize and select JERICO data from larger archives. This measure also allows to clearly define the JERICO contribution to the global marine observing system. Furthermore, procedures using the OGC’s SWE suite and SensorML format have been developed to help implement standardized descriptions of the different elements of the JERICO observing infrastructure. The descriptions can contain technical specifications of platforms and sensors, details of instrument settings, calibrations and performances, and some information on data processing procedures.

4.2. Near real-time data

Near real-time data may have several meanings:

- data that reach the data center within a few hours, and at worst, no later than 30 days, after data collection (WMO definition);
- data acquired by continuous, automatic and permanent observation networks (MyOcean);
- data that have passed an initial quality control check.

Near real-time (nRT) data can be re-submitted as delayed-mode data, particularly in the case where their original quality has been subsequently enhanced by ulterior quality checks and and/or calibrations.

4.2.1 Description of the MyOcean realtime data management infrastructure

The MyOcean data management infrastructure (Figure 3) operates through its in-situ Thematic Assembling Centres (TACs). The underlying organizational model was developed jointly with the EuroGOOS ROOSs. MyOcean is charged with gathering both nRT observations (usually within a few hours after data acquisition) for forecasting activities and aggregated historical products (1990-2012) for reanalysis.

The TAC system comprises 7 components: a central Distribution Unit (DU), hosted by the CORIOLIS network (France), and 6 in-situ TACs that, in turn, operate as DUs. The in-situ TACs and their regional domains, which can overlap sometimes, are as follows:

- SMHI (Sweden) → Baltic Sea
- IMR (Norway) → Arctic seas
- BSH (Germany) → North West shelves
- IOBAS (Bulgaria) → Black Sea
- HCMR (Greece) → Mediterranean Sea
- Puertos Del Estado (Spain) → South West shelves

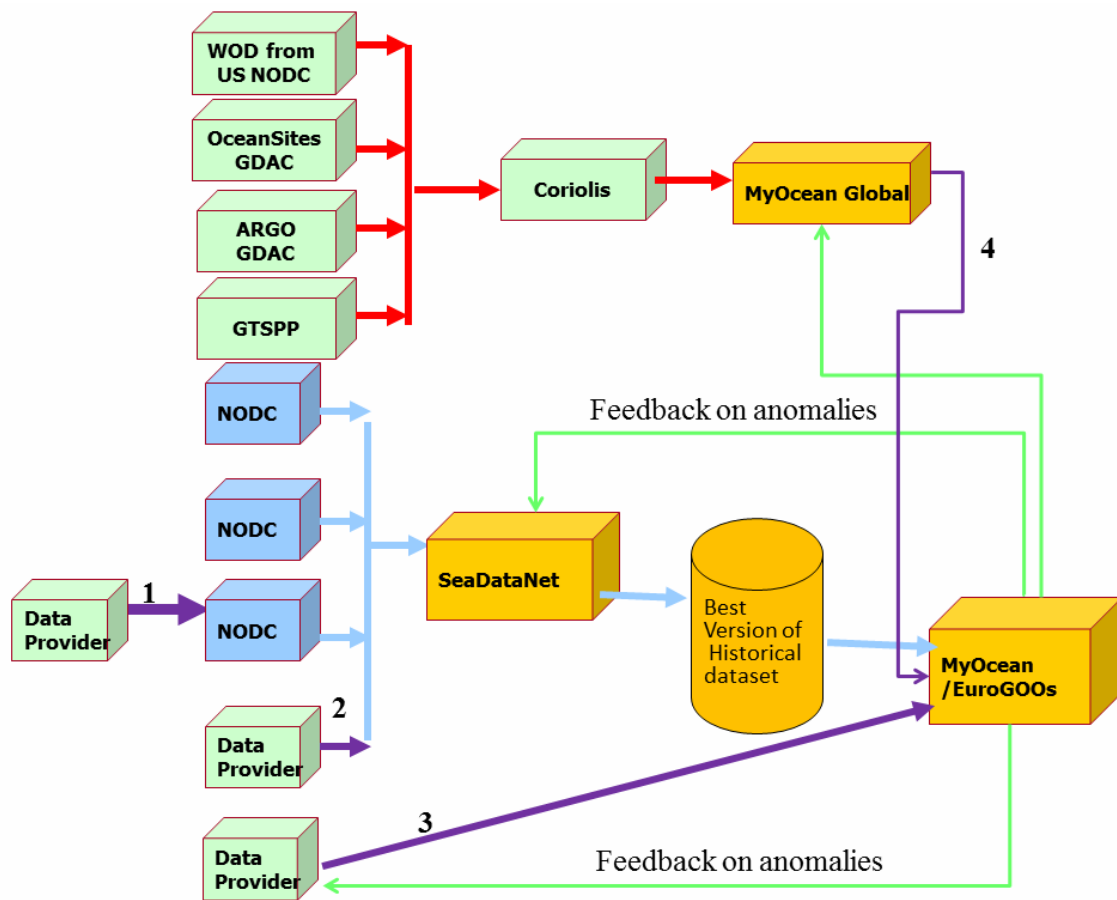


Figure 3. The main elements of the MyOcean data management infrastructure.

The central DU handles global ocean data as well as data regarding European seas. It also acts as a backup for the regional TACs. The latter serve the various ROOSs, not only in ensuring a reliable and monitored data management service but also in integrating data into useful regionally-related products.

4.2.2 Uploading the main near real-time JERICO data types

As a first step, it is recommended that new data providers contact the MyOcean TAC Help Desks or regional data managers and clarify their requirements. Generally, data transfer is effected via an FTP server. The data provider can maintain in-house data formats for any submitted data as long as these contain enough metadata information to generate the standard MyOcean NetCDF files. The information that is needed basically addresses the following concerns regarding the data: what (platform name, institution, WMO number if available, type of platform, etc.), where (location or spatial distribution), when (time or temporal interval), who (data provider, contact, PI, Data Center, etc.), and how (update or version, QC information). In some cases, it may be necessary to furnish other additional details, depending on the kind of platform involved in the gathering of the data.

4.2.2.1 Fixed platforms

Data coming from fixed platforms (moorings) should preferably be submitted using the OceanSITES format described in the OceanSITES User Manual, available at the web address, "http://www.oceansites.org/docs/oceansites_user_manual_version1.2.pdf".

In some cases, it may be possible to provide data in the MEDATLAS format (http://www.ifremer.fr/sismer/program/formats_phy/formats_UK.htm), but only when the data are being sent to HCMR, the TAC for the Mediterranean Sea.

For additional information concerning data related to fixed platforms, please visit the OceanSITES website (<http://www.oceansites.org/>).

Contacts for data submission (see Subsection 4.2.3, also):

- Regional TACs for data in NetCDF format;
- HCMR for data in MEDATLAS format.

4.2.2.2 Ferryboxes

The Ferrybox data activity in MyOcean is coordinated by NIVA (Norway). NIVA is also in charge of feeding the CORIOLIS DU with Ferrybox data to permit their widest possible dissemination.

A concerted action to set up a joint approach to manage Ferrybox data is still in the planning stages. So, at the moment, standards for uploading relative data and metadata

have not yet been compiled. Prospective Ferrybox data providers should consult the contacts indicated below for instructions on how to upload their data.

Contacts for data submission (see Subsection 4.2.3, also):

- Pierre Jacquard (NIVA);
- Beng Carlson (SMHI);
- Loic petit de la Villeon (IFREMER).

4.2.2.3 Gliders

Data coming from gliders should be submitted following the EGO ("<http://www.ego-network.org/>") recommendations developed during the FP7 GROOM ("<http://www.groom-fp7.eu/doku.php>") project.

Glider data have to be uploaded in the EGO NetCDF format (version 1.1), as described in the GROOM User Manual. The following web addresses provide details of the procedure to be followed:

- http://www.coriolis.eu.org/content/download/22398/154432/file/groom_gliders_user_manual-version1.1-sandiego.pdf;
- <http://www.coriolis.eu.org/Observing-the-Ocean/GLIDERS/EGO-gliders/EGO-Glider-data-management/How-to-provide-glider-data-to-operational-oceanography-users>;
- <http://www.coriolis.eu.org/Documentation/on-Gliders>.

The metadata relating to a glider data set must be made ready to facilitate its integration into the data file being prepared for delivery (see the section entitled “Data” of the GROOM handbook for particulars). If the data provider is unable to do this for some reason, it may be possible for the CORIOLIS DU (or a glider Data Assembly Center) to transform the relevant “.json” file to extract the required information and process the submitted data into its final form.

The minimum details concerning a submitted glider data set necessary to adequately compose its accompanying metadata is presented in the box below.

SUBMITTING A GLIDER DATA SET: REQUIRED METADATA ELEMENTS

```
"global_attributes" : {
  "platform_code_#M"      : "eudoxus",
  "date_update_#CM"      : "20140908100949",
  "institution"           : "CNRS / Division Technique Insu",
  "institution_references" : "http://www.ego-network.org
http://www.dt.insu.cnrs.fr/gliders/gliders.php",
  "contact"               : "bernardet@dt.insu.cnrs.fr",
  "principal_investigator" : "",
  "principal_investigator_email" : ""
},
"glider_sensor" : [
  {
    "SENSOR_MAKER" : "Sea-Bird Electronics",
    "SENSOR_MODEL" : "CTD 41cp",
    "SENSOR_SERIAL_NO" : "147",
    "parametersList" : [
      {
        "variable_name" :
"rawData.vars_sci_time.sci_water_pressure"
      },
      {
        "variable_name" : "rawData.vars_sci_time.sci_water_temp"
      },
      {
        "variable_name" : "rawData.vars_sci_time.sci_water_cond"
      }
    ]
  },
  {
    "SENSOR_MAKER" : "WET Labs",
    "SENSOR_MODEL" : "ECO Puck bb2fls v1",
```

SUBMITTING A GLIDER DATA SET: REQUIRED METADATA ELEMENTS (contd.)



```
"SENSOR_SERIAL_NO" : "997",
  "parametersList" : [
    {
      "variable_name" :
"rawData.vars_sci_time.sci_bbf12s_chlor_scaled"
    },
    {
      "variable_name" :
"rawData.vars_sci_time.sci_bbf12s_bb_scaled"
    },
    {
      "variable_name" :
"rawData.vars_sci_time.sci_bbf12s_cdom_scaled"
    }
  ]
},
{
  "SENSOR_MAKER" : "Aanderaa",
  "SENSOR_MODEL" : "Optode 5013",
  "SENSOR_SERIAL_NO" : "1266",
  "parametersList" : [
    {
      "variable_name" :
"rawData.vars_sci_time.sci_oxy3835_wphase_oxygen"
    }
  ]
}
],

"glider_characteristics_data" : {
  "TRANS_SYSTEM_ID" : ["300224010731840"],
  "PLATFORM_TYPE" : "Slocum",
  "PLATFORM_MAKER" : "Teledyne Webb research",
  "FIRMWARE_VERSION_NAVIGATION" : "7.10",
  "FIRMWARE_VERSION_SCIENCE" : "7.10",
```


SUBMITTING A GLIDER DATA SET: REQUIRED METADATA ELEMENTS (contd.)

```
"GLIDER_SERIAL_NO" : "136",
"LANDSTATION_SOFTWARE_VERSION" : "7.4",
"PROJECT_NAME" : "PERSEUS",
"PI_NAME" : "",
"GLIDER_OWNER": "Ecole Nationale Supérieure de Techniques Avancées, ENSTA",
"OPERATING_INSTITUTION" : "CNRS / Division Technique Insu"
},

"glider_deployment_data" : {
  "DEPLOYMENT_START_DATE" : "20130124130134",
  "DEPLOYMENT_START_LATITUDE" : 42.9909,
  "DEPLOYMENT_START_LONGITUDE" : 5.99436666667,
  "DEPLOYMENT_PLATFORM" : "",
  "DEPLOYMENT_CRUISE_ID" : "perseust02_01",
  "DEPLOYMENT_END_DATE" : "20130208130239",
  "DEPLOYMENT_END_LATITUDE" : 42.9909,
  "DEPLOYMENT_END_LONGITUDE" : 5.99436666667,
  "DEPLOYMENT_OPERATOR" : "Karim Bernardet"
}
}
```

As the number and kinds of sensors installed on a glider can change from one mission to another, a metadata file must be submitted for each uploaded data set.

Contact for data submission (see Subsection 4.2.3, also):

For glider data, the only contact is the Coriolis DU in its role as the EGO Glider Data Assembly Center (codac@ifremer.fr codac@ifremer.fr)

4.2.2.4 Fishing vessels (Fishery Observing Systems)

A concerted action to manage Fishery Observing System (FOS) data does not exist, yet. So, at the moment, standards for uploading relative data and metadata have not yet been



compiled. Prospective FOS data providers should consult the contact indicated below for instructions on how to upload their data.

Contact for data submission (see Subsection 4.2.3, also):

Loic petit de la Villeon (IFREMER; Loic.Petit.De.La.Villeon@ifremer.fr).

4.2.3 Useful contacts for near real-time data submission

GLOBAL DISTRIBUTION UNIT (Activity coordinated by CORIOLIS/France)

General Contact: codac@ifremer.fr

Principal Investigators:

- Thierry Carval (Thierry.Carval@ifremer.fr)
- Loic Petit de la Villéon (Loic.Petit.De.La.Villeon@ifremer.fr)

BALTIC SEA (Activity coordinated by SMHI/Sweden)

General Contact: myocean-service@smhi.se

Principal Investigators:

- Thomas Hammarklint (thomas.hammarklint@smhi.se)
- Anders Hartman (anders.hartman@smhi.se)

ARCTIC SEAS (Activity coordinated by IMR/Norway)

General Contact: myocean-service@imr.no

Principal Investigators:

- Henning Wehde (Henning.Wehde@imr.no)
- Sjur.Ringheim.Lid (Sjur.Ringheim.Lid@imr.no)



NORTH WEST SHELVES
(Activity coordinated by BSH/Germany)

General Contact: myocean-service@bsh.de

Principal Investigators:

- Kai Soetje (kai.soetje@bsh.de)
- Simon Jandt (Simon.Jandt@bsh.de)

BLACK SEA
(Activity coordinated by IOBAS/Bulgaria)

Contact with the centre: myocean-service@io-bas.bg

Principal Investigators:

- Veselka Marinova (marinova@io-bas.bg)
- Asen Stefanov (a.stefanov@io-bas.bg)

MEDITERRANEAN SEA
(Activity coordinated by HCMR/Greece)

Contact with the centre: myocean-service@ath.hcmr.gr

Principal Investigators:

- Leonidas Perivoliotis (lperiv@hcmr.gr)
- Antonis Chalkiopoulos (achalk@hcmr.gr)

SOUTH WEST SHELVES
(Activity coordinated by Puertos Del Estado/Spain)

Contact with the centre: myocean-service@puertos.es

Principal Investigators:

- Marta de Alfonso (mar@puertos.es)
- Fernando Manzano Muñoz (fmanzano@puertos.es)

FERRYBOX DATA
(Activity coordinated by NIVA/Norway)

Contact with the centre: myocean-service@niva.no

Principal Investigators:

- Pierre Jaccard (Pierre.Jaccard@niva.no)
- Anna Birgitta Ledang (AnnaBirgitta.Ledang@niva.no)

4.3 Downloading the main near real-time JERICO data types

The data archived in MyOcean can be downloaded through the ftp servers of either the regional TACs (for data from a specific region) or the central DU (for data from various regions). Access to the FTP servers is granted after user authentication by the MyOcean system. Users need to register in order to be able to download data. The registration procedure is simple, and is described at "<http://www.myocean.eu/web/56-user-registration-form.php>". Registered users are automatically informed of any issues affecting the MyOcean service and its products, including improvements provided by new product releases.

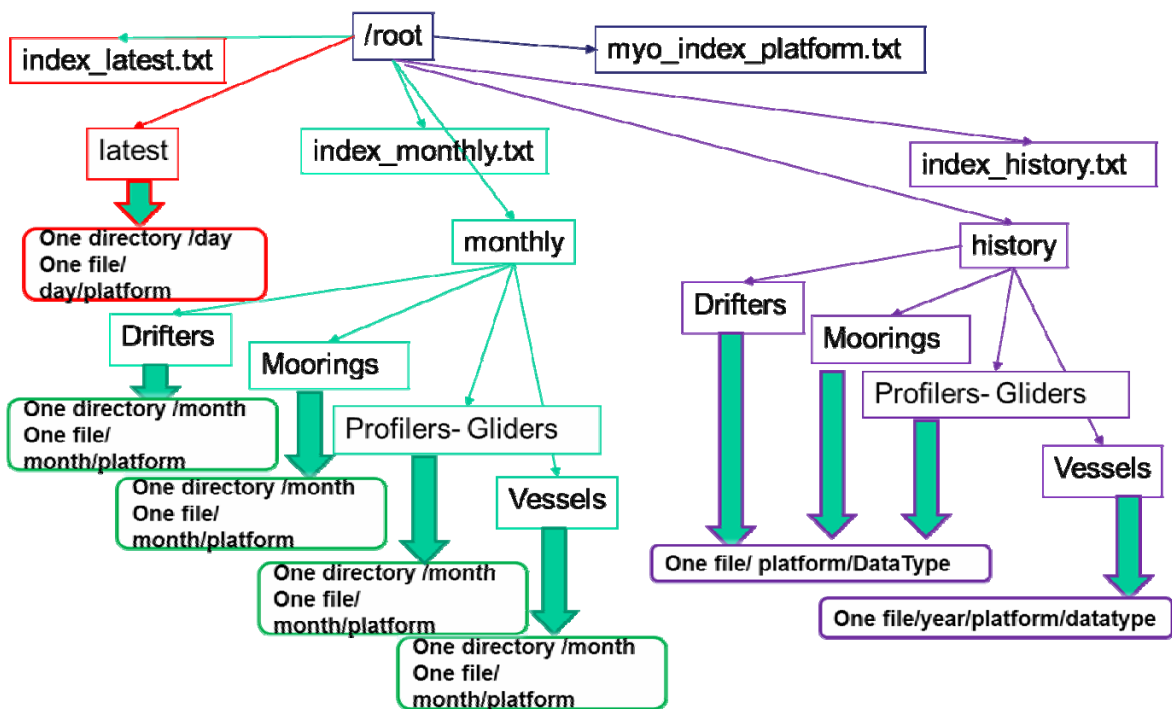


Figure 4. The MyOcean data circulation scheme



Detailed information on how to access MyOcean to get data is given in the MyOcean User Manual for In Situ Near Real Time Products (version 1.1, October 2012), available at the web address, "<http://www.coriolis.eu.org/content/download/5682/42222/file/MYO-INS-PUM-001-V1.1.pdf>". A list of relevant parameter codes can be viewed at "<http://www.coriolis.eu.org/Data-Services-Products/MyOcean-In-Situ-TAC/MyOcean-In-Situ-parameter-codes>".

The MyOcean data circulation scheme is shown in Figure 4. The organization of the in-situ TACs and the FTP servers used for distributing data is described in the MyOcean in-situ TAC documentation, available at "<http://www.coriolis.eu.org/Data-Services-Products/MyOcean-In-Situ-TAC/Documentation>".





5. Conclusions

This document gives an overview of the general data management structure and policy of the JERICO project, and furnishes instructions and information that partners involved in the project will need to join and contribute to its nRT data flow. As most of the pertinent resources, including contacts, references and useful documentation, are available online, they have been referred to by their web addresses to ensure that partners will always have access only to the most up-to-date information. The present document constitutes the second version of the JERICO Near Real-time Data Management Handbook.



6. List of acronyms

BSH: Bundesamt für Seeschifffahrt und Hydrographie.

EGO: Everyone's Gliding Observatories.

EMODnet: European Marine Observation and Data Network.

EuroGOOS: European Global Ocean Observing System.

GROOM: Gliders for Research, Ocean Observation and Management.

HCMR: Hellenic Centre for Marine Research.

IFREMER: Institut Français de Recherche pour l'Exploitation de la Mer.

IMR: Institute of Marine Research - Havforskningsinstituttet.

IOBAS: Institute of Oceanology - Bulgarian Academy of Sciences.

NIVA: Norsk Institutt for Vannforskning.

NODC: National Oceanographic Data Centre.

OGC: Open Geospatial Consortium.

ROOS: Regional Operational Oceanographic System.

SMHI: Sveriges Meteorologiska och Hydrologiska Institut.

WMO: World Meteorological Organization.

