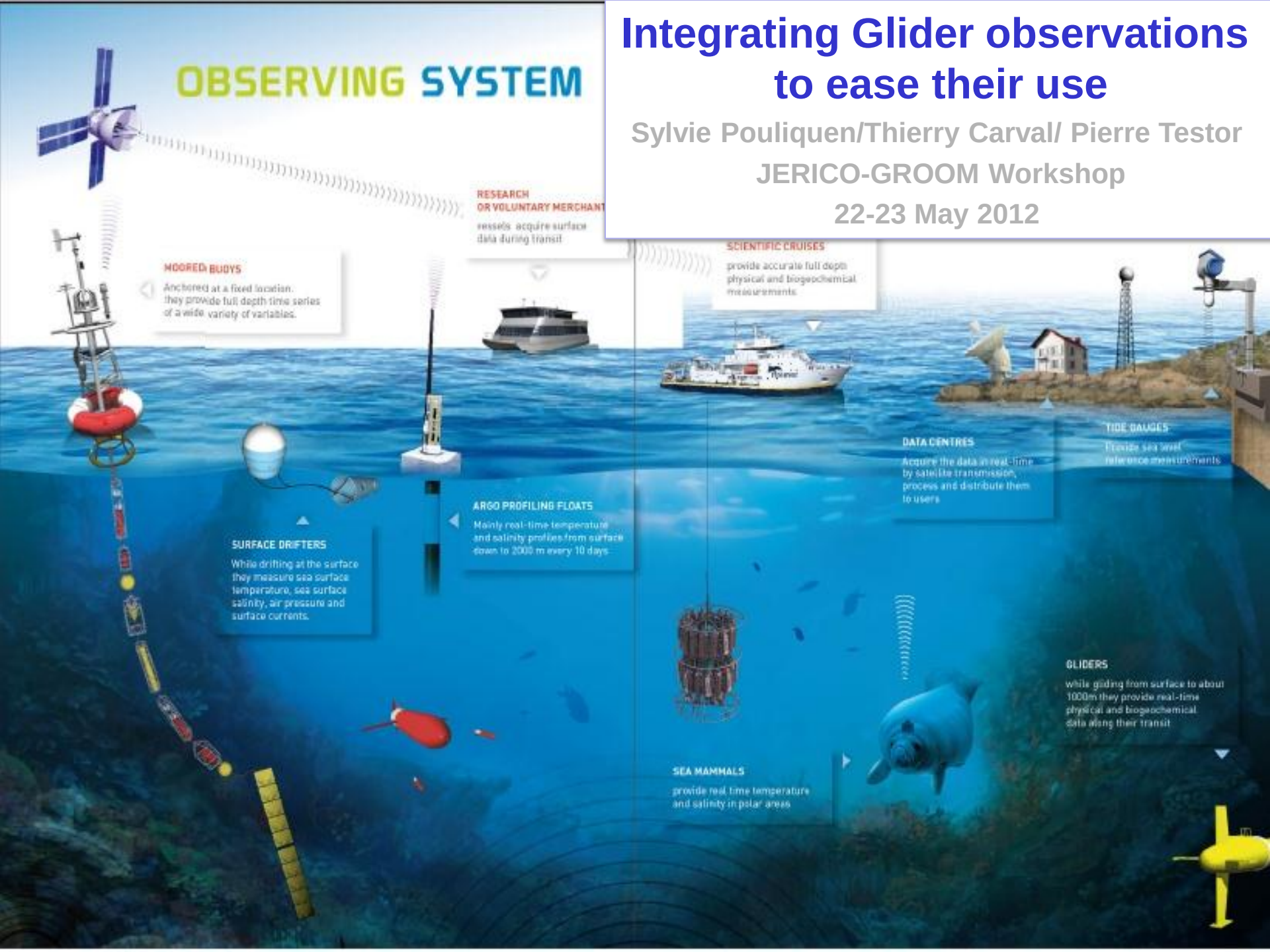
Integrating Glider observations to ease their use

Sylvie Pouliquen/Thierry Carval/ Pierre Testor

JERICO-GROOM Workshop

22-23 May 2012



What does a user expect from an observing network ?

- **Data accessible easily from a unique point**

- **Data coherent in term of :**
  - ⇒Data format
  - ⇒Data Quality
  - ⇒Processing chain (clearly documented)

- **Additional requirements for Monitoring and forecasting users**
  - ⇒Data are available in near real time (within less than 24 hours)
  - ⇒Data are available in delayed mode after calibration and /or validation

What does stakeholders expect from an observing network ?

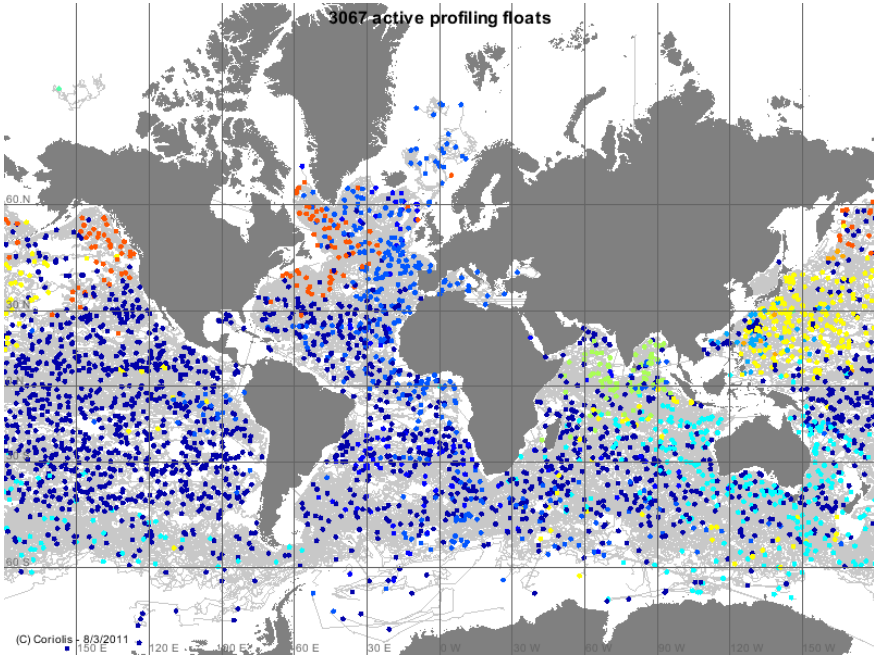
- Get more observations than they could afford alone
- Operate jointly part of the network
- Benefit from the other partners' experience from design to implementation to data management and user uptake

A key sentence for payers :

Acquire once Use multiple ?

Some success stories in the marine domain

- **Argo : more than 3000 floats sampling T &S on the global ocean from surface to 2000m on a 3°x3° grid**



•Both **Information, Scientific and Data management** activities organized since the start of Argo

•**Coordination** of the **deployments** at international level

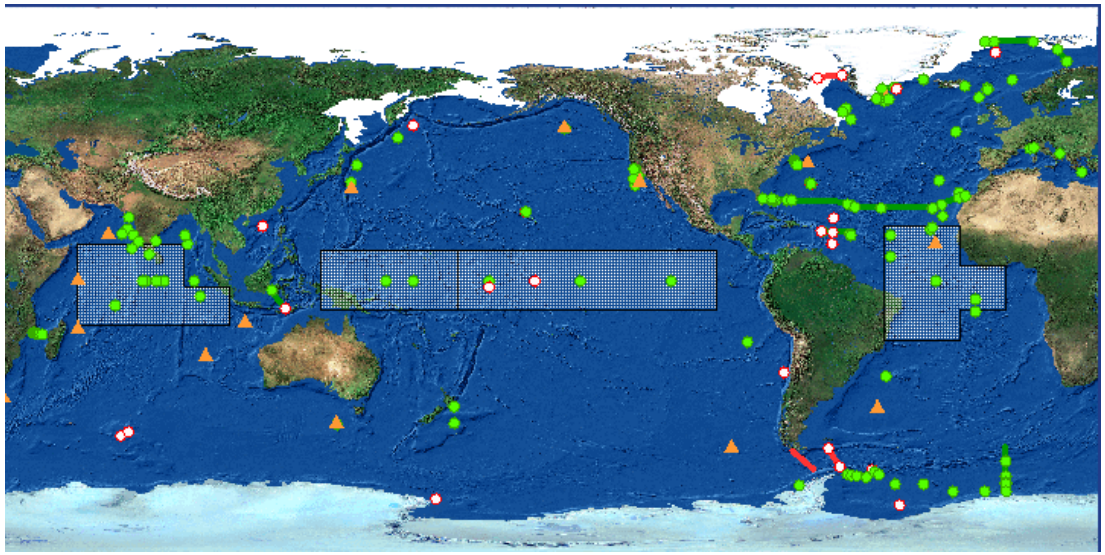
•Two single **portals** (GDACS), synchronizing themselves, to get **best copy of the Argo** data in a **unique format**

• **Common methods** for Quality control in real-time AND Delayed mode

•Organising sustainability in Europe through Euro-Argo ERIC (European Research Infrastructure Consortium)

Some success stories in the marine domain

OceanSites : 60 reference multi-parameter sites for the global ocean from surface to bottom providing long time-series



- Starting long time ago as independent research sites
- Recently set up 2 single **portals** (GDAC) to get best copy of the OceanSites data in a unique format
- **Working on Common methods for Quality control in real-time** first but also in Delayed mode

- Organising sustainability in Europe through Euro-Sites consortium

What About Gliders

- As for OceanSites , Glider activities are presently **driven by individual research drivers**
- As for Argo and some OceanSites Sites , Gliders can **deliver real-time data for core parameters** (T, S Current , Chl, O2) that are useful for both research an operational users
- Commonality with these two networks should be used for
  - ⇒**Developing integrated Data Management system**
  - ⇒**Common Data format to users**
  - ⇒**Real –time QC of core parameters**
- Gliders are **complementary to other platforms and synergy** should be developed
  - ⇒Developing a deployment strategy for other needs than pure research (ie. GMES Marine Core Service in Europe)

What is the situation today

- Data management is done by the different research communities using their own methods
- Some harmonization activities started in FP5-MFSTEP and FP6-MERSEA projects
- Data exchange in Real time is working on a best effort schema through EGO but without any commitment nor from providers or data managers
- Link with GMES MCS is done though Coriolis providing Glider data as profiles
- No agreement neither on RTQC or DMQC but best practices on RTQC through MyOcean INSTAC not widely known by the glider teams

Next Steps for Integrated Data Management of a Global Glider network

- **Organize the data management activities both**
  - ⇒at international level including the link to JCOMM
  - ⇒at regional level (EU level within GROOM and JERICO FP7, USA with IOOS, Australia with IMOS…)
  - ⇒With link with a Glider steering/science team that would provide the priorities for the network and develop the scientific procedures that need to be implemented by the Data management system

- **Set up a Glider Data Management team that would involve the Glider data managers of the Institutes operating gliders that want to be part a global Glider network**

JERICO GROOM

- **An opportunity to progress rapidly in EUROPE**
- **IN JERICO**
  - ⇒WP4 HARMONIZING OPERATION AND MAINTENANCE METHODS: calibration , bio-fouling, end-to-end Quality assessment: how to acquire good measurement at sea
  - ⇒WP5: DATA MANAGEMENT AND DISTRIBUTION : how to integrate the JERICO observation into the existing RT and Historical data system
- **IN GROOM**
  - ⇒DEFINE RT and DT data management system for Glider data in Europe

- **The projects are complementary :**
  - ⇒Should agree on the interfaces
  - ⇒Coordinate the activities for better use the funds
  - ⇒Integrate the activity into the international framework to be able to extend the network outside Europe within EGO

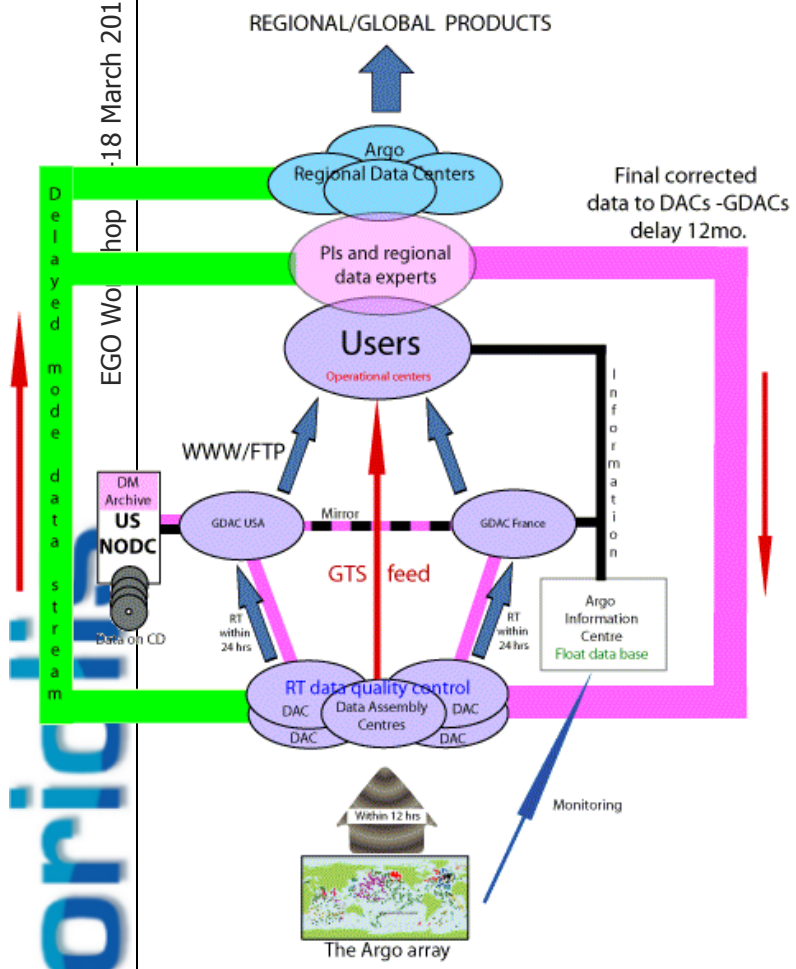
GROOM -WP3.2

Data Management for Glider

The partners

- HZG : Lucas Merckelbach , Rolf Riethmueller
- OGS : Elena Lauri
- CSIC : Tomeu Garau, Joaquin Tintore
- NURC : Reiner Onken
- PLOCAN : Carlos Barrera
- CNRS LOCEAN: Pierre Testor, Laurent Mortier
- CNRS DT-INSU : Laurent Beguery, Karim Bernardet, Elodie GONINHO
- LOV: Vincent Taillandier, Fabrizio D'Ortenzio, Lars Stemman
- HCMR : Dimitris Kassis Kostas Knittis
- NERC : David Smeed, Justin Buck, Mark Hebden
- UIB : Peter Haugan
- IFM-Geomar: Gerd Krahmann
- OC-UCY: Dan Hayes
- AWI: Agnieszka Beszczynska-Möller
- SAMS: Estelle Dumont
- UEA: Jan Kaiser, Karen Heywood
- Ifremer: S Pouliquen T Carval L Petit de la Villéon

Facilitate access to Glider data by defining a Data system for Glider data



- Based on Argo experience Define
- The Actors
- Data stream between
  - the Operators who deploy and operate the Gliders
  - the DACs who process, archive and distribute the glider data he is in charge of
  - the GDAC who integrates in a central point , in a common format, the best copy of Groom Glider data and metadata
  - The Pis , scientist that process data in delayed mode
- Data Stream should be defined in Near Real Time and Delayed Mode

Objective of task 3.2

- **Improve coherency of Glider dataset in Europe**
- **Facilitate access to Glider data by defining a Data system for Glider data**
- **The system is built to exchange glider data**
  - ⇒Must first be useful for Glider users and Glider operators and therefore provide access to as much as possible information (metadata, scientific, technical data) provided by Gliders
  - ⇒Delivering of Glider Data either on GTS or Operational users (i.e. MyOcean) will be carried on as specific delivery system that will convert the GDAC data into products usable by these operational users

How to Describe a Glider

Takes on board

Sensor S

Performs

Deployment N

Is described by

Description of the glider for this deployment

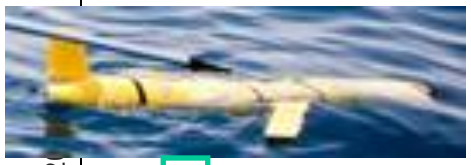
- What are the sensors on board
- What is the configuration

Transmits for each dive

Technical information measurements

Scientific measurements

Data Stream in Real time



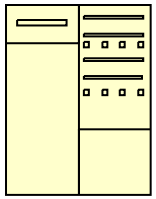
DATA

Glider Operator

DATA

DAC

Sci and Tec data



GDAC

Sci and Tec

users

Sci Data

Customize to meet other needs

GTS Users

MyOcean Users

Manufacturer Format

Common Format

Coriolis

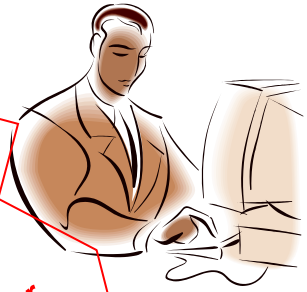
Data Stream after recovery

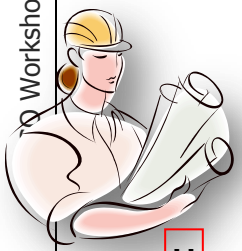
Manufacturer Format

Common Format

High Res Manufacturer Format

H R





Glider Operator

High Resolution scientific data

Real Time Data in Common format ------>
<----Real Time data procesed in delayed mode Or
Delayed mode High resolution data

H R T E C

Customize to meet other needs

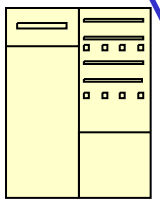
GTS Users

MyOcean Users

Sci Data

DAC

Common Format



GDAC

Common Format

users

Coriolis

Improve Coherency of Glider dataset in Europe

- **Agree on the definition of a Glider , of a deployment for a glider.**

- **Agree on common data format to share the Glider data based on OceanSites NetCDF format already used in EuroGOOS/ MyOcean/ SeaDataNet**
  - ⇒Define how to store Metadata to register the mission description and its evolution in time when changed through down link
  - ⇒Enhance OceanSites format for Scientific information by adding if necessary the new parameters sampled
  - ⇒Define how to store the technical information by definition common vocabulary
  - ⇒Study if interesting to use NetCDF4 especially for data compression

Improve Coherency of Glider dataset in Europe

- Define how to store Metadata to register the mission description and its evolution in time when changed through down link . For each deployment
  - ⇒ Static metadata : describe the sensors that are on the glider together with the calibration information
    - ➢ Define a vocabulary to describe in a common way
  - ⇒ Change in the configuration during the deployment : The operator knows the order sent to the Glider. Can he know the order that have been effectively applied by the glider?
    - The operator sends to the DAC at minimum the orders sent to the glider, better the only the orders where he is sure that they have been validated
    - The DAC update the metadata to store the history of the mission and send it to the GDAC
    - Archive at DAC level the logs of the communication with the glider. These log are not transmitted to the GDAC
  - ⇒ The changes in Glider navigation are stored as variables with the scientific data
    - Define additional parameters to register (Start from Tomeu proposal in MyOcean)

Improve Coherency of Glider dataset in Europe

- **Agree on common Near Real Time QC procedure in agreement with what exist already within EuroGOOS/MyOcean/SeaDataNet**
  - ⇒Adopt EuroGOOS/MyOcean NRT QC procedures for T&S, Chl and O_2
  - ⇒Enhance if necessary these procedures to take into account Glider specific behavior
  - ⇒Develop new recommendations in Partnership with Myo/SDN for additional parameters available in real time

How to facilitate the development of this data management system

- **Develop common tools that will transform the ascii file that are generated by the manufacturer tools into the Netcdf Common format**

- **Issues**

  ⇒Technical parameters are numerous but only a subset is transmitted in real-time and need to be shared for fleet monitoring Make a poll to identify the Realtime Technical parameters to define the mandatory technical data to be shared in real time.

  ⇒Delayed mode loop with Pi:

  - At minimum provide the Real time scientific data corrected in delayed mode
  - Ideal provide the High resolution delayed mode scientific data
  - Technical data : only the ones transmitted in real time?

Discussion