



Project of Strategic Interest NEXTDATA

Scientific Report for the reference period **01/01/2012-31/12/2012**

WP 2.2 - Archive of marine observation networks and climate reconstructions (Resp. Claudia Fratianni, INGV)

Partners: INGV

1. Scheduled activities, expected results and Milestones

The project activities for the first year are devoted to the development, realisation and organization of historical marine data archives, subjected to a specific quality control for data assimilation, and of atmospheric data that will be used in the production of Mediterranean Sea Reconstruction-Reanalysis (RR) for the past 100 years. In addition, a specific portal will be designed together with the structure of the RR database using the existing reanalysis for the past twenty years. Meetings with potential users will be organized in order to define the details of the archive and the data access protocols.

The milestone defined for the first year is:

M2.2.1 (PM12): Meeting to discuss the specific RR portal and specific products. The meeting was held in December at INGV in Bologna and it was attended by the private Company CLU s.r.l., which will be in charge of the design and architecture of the specific portal for the future.

2. Deliverables expected for the reference period

The Deliverable due for the first year is:

D2.2.1 (PM12): Report on specific in-situ and satellite data for RR and atmospheric parameters.

D2.2.2 (PM12): Report on user-friendly catalogue and database.

3. Activities which has been actually conducted during the reference period

Most of the planned activity was carried out during the first year. Specifically, a reference database was created for the historical marine data and for the atmospheric data required for the production of the RR, starting from the information already collected for the production of the previous reanalysis versions, as shown in Table 1, and proceeding in the search for new and updated datasets to cover the time interval of the Reconstruction/Reanalysis (RR).

3.1 Research activity

The activity was conducted in close collaboration with WP1.3 which, wherever necessary, developed specific data quality controls, as described in the Deliverables D1.3.1 and D1.3.2. This was done to create a reference database which should be broad enough to significantly contribute to increasing the quality of the RR.

	MedReanV2 (1985-2007)	MedReanV4(1985-2010)	RR(1912-2011)
Initial conditions	MedAtlas Climatology (obs 1995-1999) (Maillard et al. 2005)	SDNV2aa climatology (obs until1987)	SDNV2aa climatology (obs until1987)
Atmospheric forcings	ERA15 1.125° (1985-1992) ECMWF analyses 0.5° (1993-2007)	ERAInterim 0.75° (1985-2010)	AMIP1.125° (1912-1957) ERA40 1.125° (1958-1978) ERAInterim 0.75° (1979-2011)
Total Cloud Cover	NCEP-NCAR (1985-92) ECMWF analyses 0.5° (1993-2007)	ERAInterim 0.75° (1985-2010)	AMIP 1.125° (1912-1947) NCEP-NCAR(1948-1978) ERAInterim 0.75°(1979-2011)
Precipitations	NCEP-NCAR (monthly climatology)	CMAP (monthly climatology)	AMIP (monthly climatology) NCEP-NCAR (monthly climatology) CMAP (monthly climatology)
SST	SST reconstruction (1985-2007) (Marullo et al.2007)	SST reconstruction (1985-2007) (Marullo et al.2007) DT data (2008-2010)	HadISST (1912-1985) SST reconstruction (1985-2007) (Marullo et al.2007) MyOcean data (2008-2011)
SLA	ERS1, ERS2, EnviSat, TOPEX/Poseidon, Jason1 (Pujol and Larnicol 2005)	AVISO multisensor "UPD" data reprocessed in 2010	MyOcean multisensor "UPD" data reprocessed in 2012
XBT	<1999 MEDATLAS, MATER Ship of Opportunity (Manzella 2007) 2000-2007 MFS observations	SeaDataNet MFS observations (1985-2010)	SeaDataNet MFS - MyOcean observations (1912-2011)
ARGO	MedArgo Program 2001-2007 (Poulain et al. 2007)	All observations 2001-2010	All observations 2001-2011
CTD	MEDATLAS, MATER (<1999) 2000-2007 MFS observations (2000-2007)	SeaDataNet MFS observations (1985-2010)	SeaDataNet MFS - MyOcean observations (1912-2011)

Table 1. Table of datasets used for the production of reconstructions and reanalyses at INGV. The systems are described in Deliverable D1.3.1.

The datasets stored for the production of the RR are as follows:

- satellite sea surface temperature (SST) observations;

- satellite sea level anomaly (SLA) observations;
- in situ temperature and salinity observations;
- surface atmospheric variables;
- temperature and salinity initial conditions.

SST observations:

The creation of a SST dataset covering the whole RR period focused mainly on the search for available data for the period before 1985. After this date, in fact, there exist specific satellite SST products created for the Mediterranean Sea, consisting of daily time series of mean sea surface temperature maps optimally interpolated onto the RR model grid at 1/16th of a degree. As from 1985, the dataset is a result of the concatenation of the following products:

- 1985-2008.07.10: reconstruction data built from the most recent AVHRR Pathfinder SST time series (Marullo et al., 2007).
- 2008.07.11-2011.12.31: in the framework of the MyOcean project, SST TAC/OSI TAC (Sea Surface Temperature/Ocean and Sea Ice Thematic Assembly Center)) produce and disseminate SST daily maps, as described in Deliverable D2.2.1.

The Met Office Hadley Centre SST dataset (HadSST1) was archived for the period preceding 1985. It consists of monthly SST data on a regular grid of 1° x 1° starting from 1870 (Rayner et al., 2003). This choice is consistent with the idea to use AMIP (Atmospheric Model Intercomparison Project – Gates, 1992) type of atmospheric forcing for the pre-ERA40 period (which starts from mid-1957). AMIP data are obtained from ECHAM4 model simulations forced by HadSST1.

Satellite SLA observations:

The SLA dataset was updated to the latest version, completing the time series till April 2012. These data are produced and disseminated in the framework of Sea Level TAC of the MyOcean Project.

The dataset consists of mono altimeter satellite along-track sea surface height anomalies computed as difference between sea surface height and the mean over the period 1993-1999. (<http://catalogue.myocean.eu.org/static/resources/myocean/pum/MY02-SL-PUM-008-001-005-v3.2.pdf>).

The coverage depends on the duration of the missions:

- Jason 2: from October 2008
- Jason 1 (new orbit): from February 2009
- Jason 1: April 2002 - October 2008
- Envisat (new orbit): from October 2010
- Envisat: October 2002 - October 2010
- ERS-1: October 1992 - May 1995
- ERS-2: May 1995 - April 2003
- GFO: January 2000 - September 2008

- T/P (new orbit): September 2002 - October 2005
- T/P: September 1992 - April 2002

Each satellite is characterised by a repeat cycle and by a ground track which, overall, guarantee relatively uniform coverage of the Mediterranean Sea.

The information stored for each satellite is that required by the assimilation system, i.e.: time in Julian days starting from 1 January 1950, longitude, latitude, and sea level anomaly value expressed in meters.

In situ observations of temperature and salinity vertical profiles

The dataset is the result of the homogenization of data collected in various projects, namely MedarMedatlas, SeaDataNet, MFS and MyOcean in order to eliminate duplicates. The dataset is composed of vertical temperature and salinity profiles collected by means of bottles, thermometers, XBT, MBT, CTD and ARGO measurements subjected to a specific quality control for use in the data assimilation system, as described in the Deliverable D1.3.1. The composition of the reference dataset for the in situ observations is described in the Deliverable D2.2.1.

Surface atmospheric variables

The parameters required to force the RR system are:

1. Mean sea level pressure (MSLP)
2. Temperature at 2 meters (T2m)
3. Zonal and meridional wind component (U10,V10)
4. Total cloud cover (TCC)
5. Dew point temperature at 2 meters (D2m)
6. Precipitation (P)

The creation of the reference dataset calls for the concatenation of different starting datasets in order to ensure the time coverage of the RR, as shown in Table 1. The parameters (MSLP, U, V, TCC, T2m, and D2m) were retrieved from the reference datasets and stored in monthly or daily binary files which will be read by the oceanic model and which will contain the variables at the spatial and temporal resolution of the original model. The precipitation parameter was archived as a NetCDF file containing the monthly precipitation values interpolated from the original grid to the oceanic model grid at 1/16 of a degree.

The original datasets (AMIP, NCEP-NCAR, ERA-40, ERAINTERIM, and CMAP) are described in the Deliverable D2.2.1.

Temperature and salinity gridded fields (Initial Conditions)

The initial conditions are stored in the form of two separate files in NetCDF format, one for salinity and one for temperature. The method utilised to create the files is described in the Deliverable D1.3.1.

3.2 Applications; technological and computational aspects

After checking, the data are stored in a dedicated filesystem as shown in Figure 1.

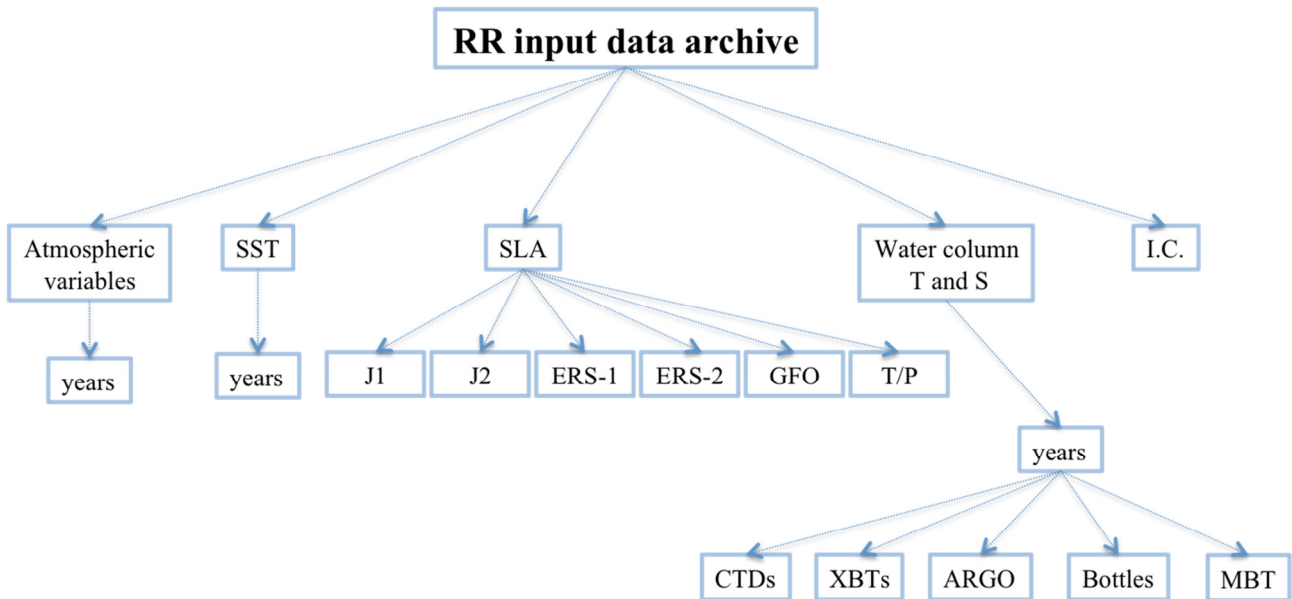


Figure 1. Organisation of RR system input data archive directories.

All datasets, divided by years and/or instruments, are written in a format which can be read by the OceanVar data assimilation scheme and the numerical model:

- ASCII: T and S observed profiles; SLA satellite observations;
- binary: surface atmospheric forcing;
- netCDF: SST observations; T,S gridded fields; precipitation.

The specific RR data access portal was designed in parallel with the implementation of the reference database. In the creation of the specific RR portal, the most recent guidelines of the European Community INSPIRE Directive were followed in order to create an infrastructure that is compatible with the European Community framework. The specific portal will be linked to the General Portal of the project, which will contain all information on the metadata of the products that will be made available. The property and network services will reside on the specific thematic portal, as shown in Figure 2.

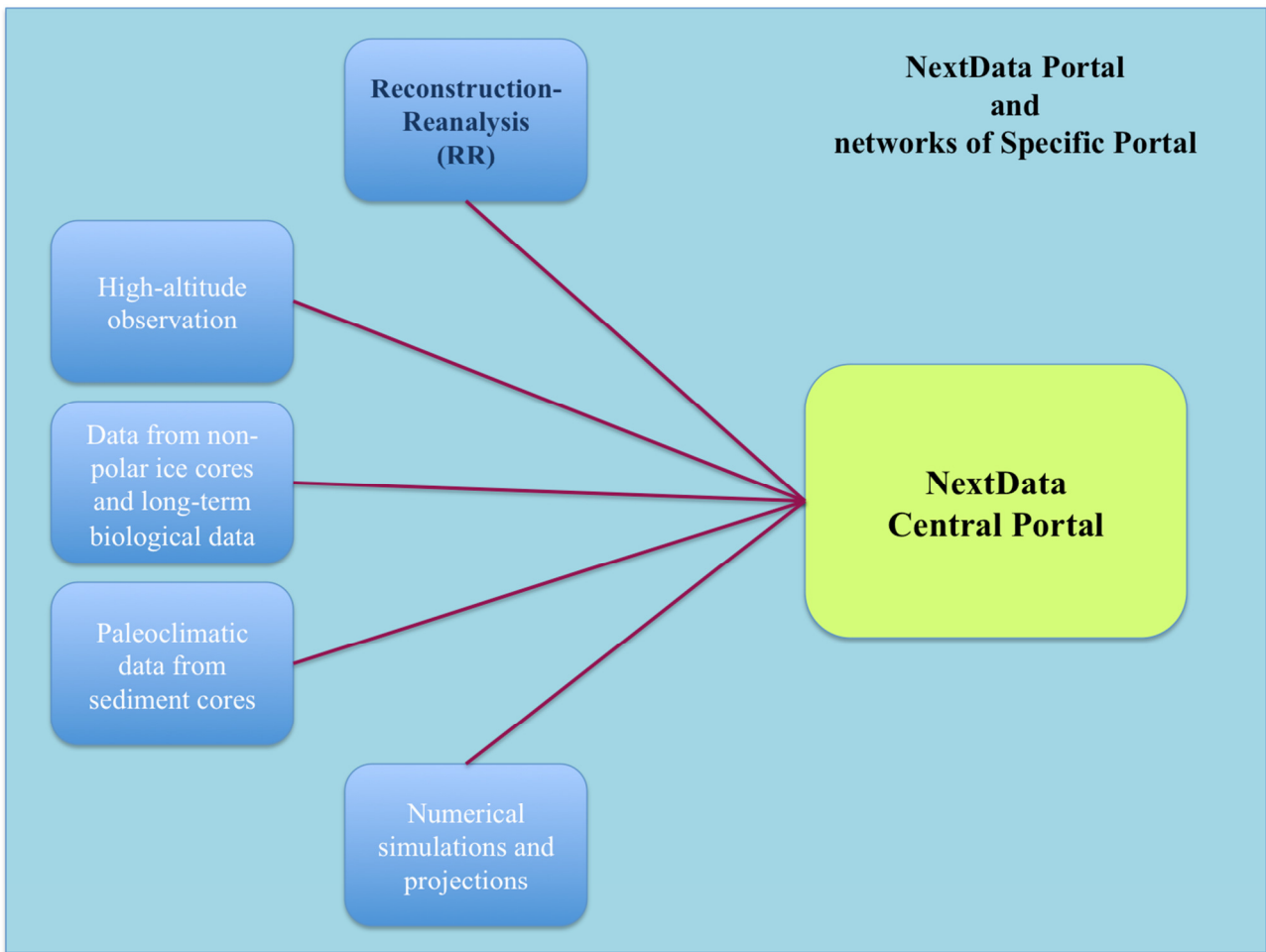


Figure 2. Diagram of the information flow between the General Portal and the specific thematic portal of the RR.

The specific thematic portal will supply the services defined in the INSPIRE Directive: discovery, view, downloading and transformation, as shown in Figure 3.

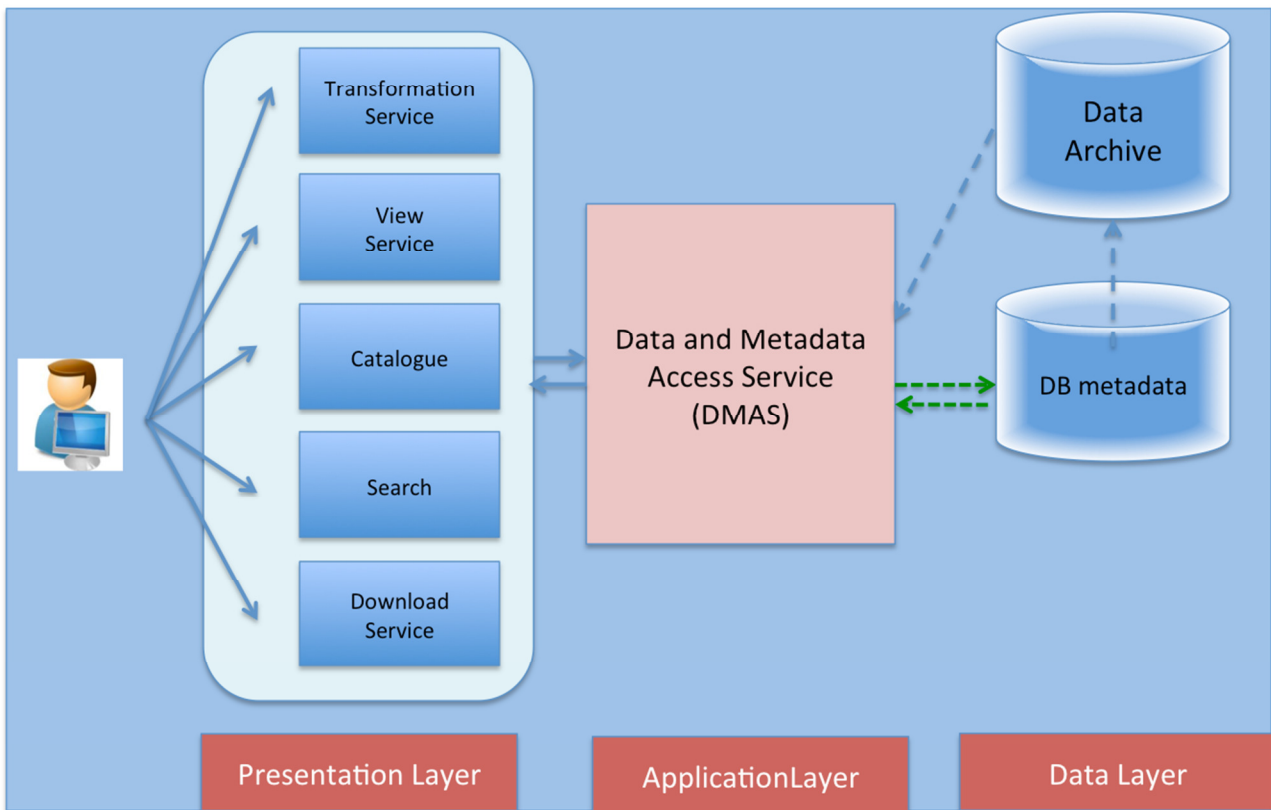


Figure 3. Architecture of the specific RR thematic portal.

The specific RR thematic portal architecture is composed of 3 closely interconnected layers:

- **Presentation Layer:** this layer contains the network services (Transformation, View, Search, Download, Catalogue) that it was decided to implement.
- **Application Layer:** includes the Data and Metadata Access Service (DMAS).
- **Data Layer:** includes the physical archive of the data and the metadata.

The user will access a graphic interface (web page) that will provide a simple method of consulting the catalogue, searching, viewing and downloading RR products.

The heart of the entire architecture is constituted by the DMAS, which performs a dual function. On the one hand, it manages the metadata required to perform search and discovery functions in the data archive, and on the other hand it provides the facility to access the data archive by means of the metadata database. The data archive is the physical storage in which the products will be located, and it will consist of a filesystem. This solution was adopted in view of the enormous quantity of data to be handled and the need for a tool that is separate from database management systems.

The user can consult the catalogue of available products, which will include the parameters indicated in Table 2. This catalogue was designed on the basis of the currently produced reanalyses, which cover the last 20 years, and it may be subject to updates in versions that will be released subsequently during the project.

Variables	Geographical coverage	Spatial resolution	Number of levels	Temporal resolution	Temporal coverage
Temperature	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	33	24 hr average field	1912-2011
Salinity	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	33	24 hr average field	1912-2011
Sea Surface Height	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	1	24 hr average field	1912-2011
Horizontal velocity (meridional and zonal component)	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	33	24 hr average field	1912-2011
Surface fluxes	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	1	6hr	1912-2011
Surface S,T and currents	6° W - 36.25° E 30.19° N – 45.94° N	0.0625°	1	6hr	1912-2011

Table 2. List of the RR products made available in the framework of the project.

Temperature, salinity and horizontal velocity fields:

These are 3-dimensional fields, given as daily means centered at 24 UTC of each day. The vertical levels are in meters and the geographical coordinates are in degrees and tenths of a degree. A study will also be performed to determine how to filter the signal from inertial currents by means of a high-pass filter.

Surface horizontal temperature, salinity, and velocity fields:

These are 2-dimensional fields, given as instantaneous fields (snapshots) every 6 hours.

Free surface elevation fields:

These are the 2-dimensional fields, given as daily means centred at 24 UTC of each day.

Air-sea interface fluxes fields:

These are 2-dimensional fields, given as instantaneous fields (snapshots) every 6 hours. They include the heat fluxes with all their components (net short wave radiation, net long wave radiation, sensible heat and latent heat), momentum (zonal and meridional wind stress) and surface water flux (difference between precipitation, run-off and evaporation).

The RR products are released on the regular lat/lon grid of 1/16 of a degree, equivalent to approximately 6.5 km. The reference domain is shown in Figure 4.

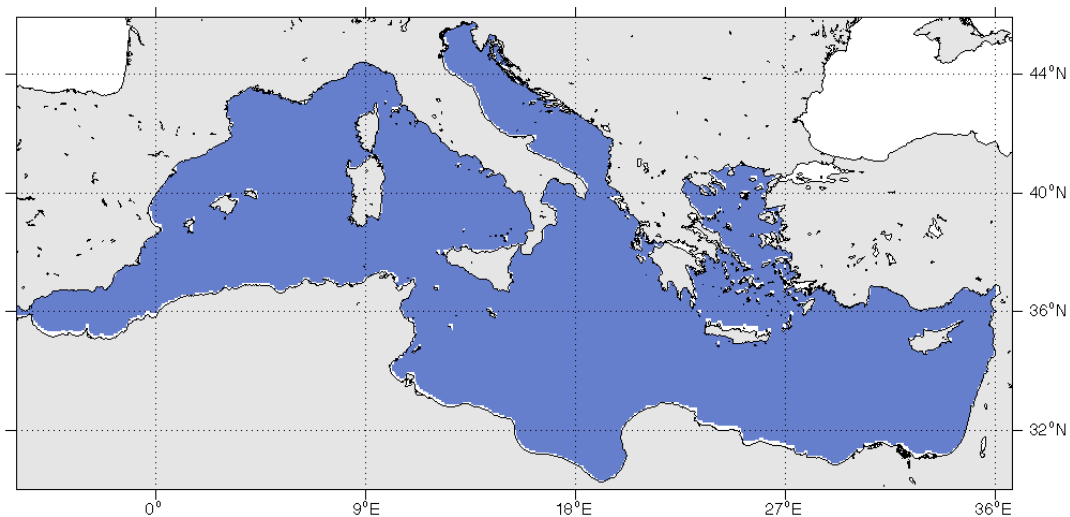


Figure 2. RR domain.

The three-dimensional parameters will be made available on 33 levels corresponding to the IODE standard levels.

The Portal interactive viewing options will be possible only for mean fields such as:

- Monthly mean
- Annual mean

This choice is justified by the large volume of daily data that will not allow to have good performances in terms of request/response time of the Portal. So interactive visualization will be allowed only for mean fields.

It will be possible to download the selected products in NetCDF (Network Common Data Form) format, a binary format supported by specific libraries and widely used in the scientific field (www.unidata.ucar.edu/software/netcdf/). The convention adopted is CF-1.0.

3.3 Formation

None in the reference period.

3.4 Dissemination

None in the reference period.

3.5 Participation in conferences, workshops and meetings

None in the reference period.

4. Results obtained during the reference period

4.1 Specific results (database, measurement results, models output, etc.)

The reference database of historical marine data and atmospheric data was created for use in the production of the RR, as described in section 3.

4.2 Publications

The completion of publications during the reference period was not envisaged.

4.3 Availability of data and models output (format, media, etc.)

No products planned in the reference period.

4.4 Completed deliverables

Both deliverables envisaged for the first year were completed.

5. Comment on differences between expected activities/results/deliverables and those which have been actually realised

The meetings with potential users to define the details of archive and data access protocols will be postponed to the second year. First contacts are taken with the WP involved in archiving and releasing of reanalysis and simulation gridded products in order to find a common management strategy.

6. Activities planned for the next period

In the second year of the project, the discussion on the archive organization and on the products release will be opened to the scientific community. The historical marine data archive will be updated following the dataset updates. In the second year of the project, the archive will be created with the produced RR data and a second version of the catalogue will be implemented on the basis of the new products. In addition, the interactive system for viewing and downloading the archive will be developed. On the base of the obtained results, it will be evaluated when to organize a meeting with the community of the end users, with the release of the beta archive version and the specific thematic portal.