Mediterranean Marine Science

Vol 12, No 3 (2011)

Vol 12, No 3 (2011) special issue

Data Centre for French Coastal Operational Oceanography

M. FICHAUT, A. BONNAT, T. CARVAL, F. LECORNU, J. F. LE ROUX, E. MOUSSAT, L. NONNOTTE, S. TAROT

doi: 10.12681/mms.72

Mediterranean Marine Science
Indexed in WoS (Web of Science, ISI Thomson) and SCOPUS
The journal is available on line at http://www.medit-mar-sc.net

Paper presented at the
"International Conference on Marine Data and Information Systems - IMDIS", Athens, 2008

Data Centre for French Coastal Operational Oceanography

M. FICHAUT [1], A. BONNAT [1], T. CARVAL [2], F. LECORNU [3], J.F. LE ROUX [3], E. MOUSSAT [1], L. NONNOTTE [4] and S. TAROT [4]

[1] IFREMER, IDM/SISMER, BP 70, 29280 Plouzané, France
[2] IFREMER, IDM/ISI, BP 70, 29280 Plouzané, France
[3] IFREMER, DYNECO/PHYSED, BP 70, 29280 Plouzané, France
[4] Altran Ouest, Technopôle Brest Iroise, 300, rue Pierre Rivoalon, 29200 Brest, France

Corresponsing author: Michele.Fichaut@ifremer.fr

Abstract

The data centre for French coastal operational oceanography (CDOCO) has been developed in the frame of PREVIMER, the coastal operational system for the French marine environment. The main objective of a coastal oceanography data centre is to provide in a timely fashion the data needed by modellers for assimilation, forcing and validation purposes. The goals of this data centre are:

- to provide operational services to the actors and users of coastal operational oceanography,
- to develop partnerships with the data producers and the data users,
- to develop the operational services for data collection, quality checks, archiving and distribution.

The data are received as much as possible in real time or near real time and the data centre delivers it to the modelling processes for several daily runs. In order to feed the models, the data centre collects:

- in situ oceanographic measurements,
- meteorological data (wind, temperature, atmospheric pressure, rainfalls, radiation),
- hydrology information (rivers outflows and nutrient fluxes),
- limit conditions at open boundaries (Mercator, MFS, climatologies),
- bathymetry.

Satellites observations are also taken in consideration but they are processed by the CERSAT data centre and do not circulate through the CDOCO. Two different data flows can be distinguished: a real time data flow which manages the data delivered by the partners (SHOM, Météo-France, DIREN and IUEM) and a delayed mode data flow which manages other data such as reference data (bathymetry); climatologies, analyses.The data centre automatically collects all these data that are received at different frequencies (from several times a day to once a week). Both observed data collected for the models and model results (hindcast and nowcast analyses and forecasts) are archived. It is also of the responsibility of CDOCO to distribute (under agreement) the data and results in a timely fashion.The data are distributed on a timely basis to the actors of the PREVIMER operational prediction system. A password protected access is also provided to other users for downloading data through web interfaces.The data are distributed using standardized specific formats: NetCDF and ASCII.

Keywords: Coastal oceanography; Data management; Integrated database; PREVIMER.

Introduction

The main objective of a coastal operational oceanography prediction system is to provide an efficient means to evaluate the state of the environment for the purposes of security of the activities and management of resources.

Coastal models run on a daily basis and produce real time forecasts. To be able to run, these models need several kinds of data. It includes data as different as ocean in situ data, atmospheric data, river flow data and other model results.

This paper presents in details the different functionalities that have been developed within the coastal operational oceanography data centre (CDOCO) to provide an efficient operational data service. It also details all the data that are processed (from data collection to final distribution) in the CDOCO and the system developed to monitor the data collection and distribution.

Context and objectives of the data centre

In the field of operational coastal oceanography, PREVIMER is a French project which aims to provide users of coastal oceanography with observations, modelling tools and real time forecasts (LECORNU, 2008). Its goals are to provide 48 hours to 6 days coastal forecasts and observations for the French coastlines (i.e. in the British Channel, the Bay of Biscay and the NW Mediterranean Sea) for the following parameters: direction and intensity of currents, temperature, salinity, sea level, waves (frequency, direction and height), concentration of particles or plankton and water quality.

PREVIMER addresses general public in recreational activities (surfing, swimming, diving, yachting…), professionals (shell farms, fisheries, maritime traffic and off-shore industry), and local authorities managing the coastal environment, scientists (European and international cooperation) and consultants impact studies.

To provide an operational service PREVIMER relies on specific technical infrastructures such as the Regional Performance Computer System based on the powerful CAPARMOR super computer (ODAKA, 2008), the hardware infrastructure for the data archiving and the already existing databases managed in the Ifremer's department "*Informatique et Données Marines (IDM)*". The data centre for French coastal operational oceanography (CDOCO) is one of these technical infrastructures which were developed to support PREVIMER which is the main user of CDOCO.

The PREVIMER system is based on pilot numerical models (called demonstrators) which concern several geographical areas displayed on Figure 1.

The models cover several themes in physical and biological oceanography; they focus on the following parameters:

- temperature, salinity, sea level, currents direction and intensity and waves frequency, direction and height in the Iroise Sea (Demonstrator 1, D1),
- primary production over the Bay of Biscay (D3),
- turbidity in the Bay of Vilaine and Southern Brittany (D4),
- primary production, nutrients and plankton concentrations over Brittany (D6),
- temperature, salinity, currents and seal level over the French seafronts (F1-F2).

PREVIMER runs several models on a regular basis:

- MARS (LAZURE, 2008) developed in IFREMER for temperature, salinity, currents, sea level and primary production,
- WaveWatch III, REF/DIF and SWAN for waves.

In this context, the tasks of the French coastal operational oceanography data centre are:

- to provide operational data delivering services to the actors and users of coastal operational oceanography who need information on currents, sea level anomaly forecasts, monitoring of sea water quality and urban contaminants discharge, or on phytoplankton blooms,
- to develop partnerships with the data producers and the data users,

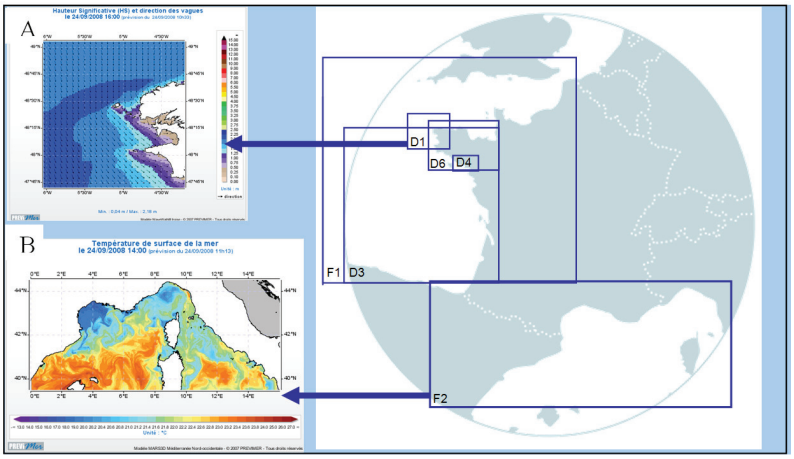

Fig. 1: Main geographical areas of PREVIMER models and some results: A: Waves in the Iroise Sea (D1), B: temperature in the western Mediterranean Sea (F2).

- to develop the operational services for data collection, quality checks of the in situ data, archiving of the collected data and of the model results and distribution of the data to the final users.

Setting up the data centre

As the coastal models need a wide variety of data types to be able to produce synoptic observations and 48 hours to 6 days forecasts on the French marine coastal areas and because the models are running daily in operational mode (or twice a day for the wave models), 7 days in 7, setting up a data centre for coastal operational oceanography was quite a challenge.

Ifremer and the partners of the Coriolis project already had a confirmed expertise in managing in situ data for the needs of operational oceanography. When building the CDOCO infrastructure, it was decided to rely, each time it was possible, on what was developed for the CORIOLIS DATA CENTRE. Coriolis plays an important role in international projects (Global Data Centre for ARGO -T & S vertical profiles-, Surface Underway data -GOSUD-, Moorings data -OceanSites-). The Coriolis data centre is the data provider for MERCATOR, MERSEA and SOPRANE forecast systems. Coriolis is also able to manage delayed mode data set and reference global datasets for validation purposes.

However, in the case of PREVIMER new parameters have to be taken into account (waves, chlorophyll, currents and meteorological) and also data for model forcing which are specific to coastal areas such as river outflows and precipitation rates. The creation of the CDOCO, specific to coastal data, but using existing infrastructure, was necessary.

This creation was conducted in four stages. The first step was to define the list of data necessary to PREVIMER demonstrators. Then for some of these data it was necessary to develop partnerships and agreements with the data providers. The third step consisted in the software development and in setting up the hardware infrastructure (discs for data archiving), and finally the operational data management structure had to be defined.

Data inventory and data collection

The models need to integrate several types of data to be able to deliver coastal forecasts: forcing and reference data, in situ and spatial measurements.

Table 1
Model outputs used as forcing data collected by CDOCO.

Forcing data	Domain	Model originator	Data provider	Collection frequency
ARPEGE	Wind and thermal fluxes	Météo-France	Météo-France	7 times / day
ALADIN				7 times / day
GFS High Resolution [1°]		US National Weather Service	NCEP	4 times / day
GFS High Resolution [0,5°]				4 times / day
NCEP Brittany/Gascogne (NCEP WW3)				4 times / day
Mercator-Atlantic ocean		MERCATOR	MERCATOR	Once / week
MFSTEP		INGV (Italy)	INGV	Once / day
MM5		Pennsylvania State University/NCAR	ACRI	Once / day
WaveWatch3 Atlantic	Waves	NOAA	SHOM	Once / day
WaveWatch3 Biscay				Once / day
WaveWatch3 Global				Once / day
WaveWatch3 Mediterranean Sea				Once / day
WaveWatch3 Spectra Biscay and Mediterranean Sea				Once / day
GR4J	Freshwater input	CEMAGREF	IFREMER/ DYNECO	Once / day

Forcing data

The forcing data (Tables 1 and 2) can be divided into three main categories:

PREVIMER and research laboratories models need also data on sea water/atmosphere interface such as wind and thermal fluxes. These data are data produced by PREVIMER partners like numerical weather prediction with the models ARPEGE and ALADIN developed by Météo-France, or public data like NCEP produced by the US National Weather Service and MM5 produced by Pennsylvania State University.

Forcing factors for the coastal models must also include freshwater input such as river flows and precipitation rates, which are in situ measurements received in near real time from different sources. When river flows data are not available, which is the case for most small rivers, IFREMER uses the GR4J model output to compute daily the river flows in Brittany. CDOCO collects also these data and model results.

Results of global models running at the basin scales are used as limit conditions at open boundaries, by PREVIMER models. These data are MERCATOR forecasts and hindcasts for the Atlantic Ocean, MFS for the Mediterranean Sea, and swell fields from WAVEWATCH III.

In situ measurements

In situ measurements are essential for validation of the models, to be compared to the model's results. It is necessary to get as many data as possible in real time, but data received in delayed mode are also valuable for *a posteriori* validation.

The data are received as much as possible in real time or near real time and the data centre delivers it to the modelling centre for the daily runs (Fig. 2).

CDOCO collects in situ data from several measurement networks (Table 3):

- MAREL fixed buoys network: automated

Table 2
In situ data used as forcing data collected by CDOCO.

Data type	River or station	Data provider	Collection frequency
River flows	Adour	DIREN Aquitaine (SP Coliane)	4 times / day
	Dordogne	DIREN Aquitaine (SP Coliane)	4 times / day
	Garonne	DIREN Midi-Pyrénées (SP Coliane)	4 times / day
	Gave de Pau	DIREN Aquitaine (SP Coliane)	4 times / day
	Gave d'Oléron	DIREN Aquitaine (SP Coliane)	4 times / day
	Loire	DIREN Pays de la Loire (SP Coliane)	4 times / day
	Luy	DIREN Aquitaine (SP Coliane)	4 times / day
	Nive	DIREN Aquitaine (SP Coliane)	4 times / day
	Rhône	Compagnie Nationale du Rhône (CNR)	2 times / day
	Seine	DIREN Ile-de-France	Once / day
	Vilaine	Institut de l'Aménagement de la Vilaine (IAV)	Once / day
Precipitation rate	Guipavas	Météo-France	9 times / day
Potential evapotranspiration	Brittany	Météo-France	Once / day

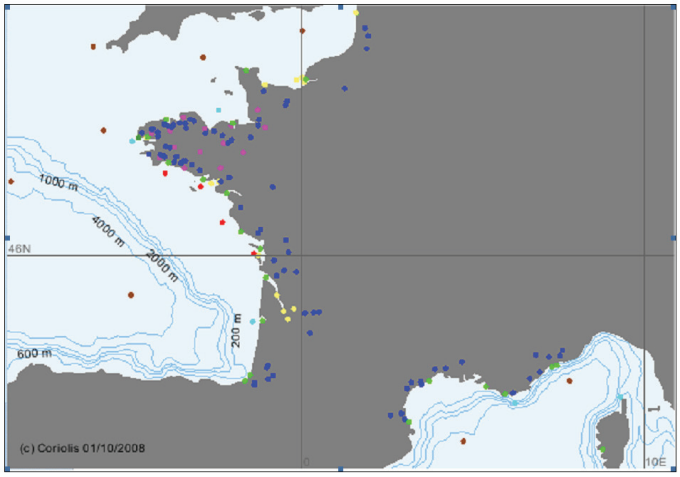

Fig. 2: Map of the in situ fixed station data, collected in CDOCO (including river flows, meteorological. stations on land and at sea, islands network sensors, tide gauges, swell buoys, MAREL buoys).

acquisition and transmission of data related to the coastal water quality.

- RECOPESCA network: data-recording network of fishing activity and environmental parameters which involves volunteer participation of fishermen. The environmental data are temperature and salinity, collected by CTD sensors installed on trawls and traps of the fishing boats and transmitted to CDOCO in near real time through on-board secured transmission systems.
- Sea level data from the tide gauge network of the French Hydrological Service (SHOM).
- Swell buoys network of CETMEF in charge

Table 3
In situ data collected by CDOCO.

Sensor type	Data	Data provider	Frequency	Quality Checks by
Buoys	Marel-Iroise	IUEM	Once / day	data provider
	Marel-ROSLIT	IFREMER	Once / day	data provider
	Marel-Gironde	University of Bordeaux-1/ EPOC	Once / day	data provider
	Drifting buoys	IFREMER	2 times / day	by Coriolis
	Meteorological buoys (IAA)	Météo-France	Once / day	data provider
	Swell buoys	CETMEF	6 times / day	CDOCO (automatic and visual)
	MOLIT	IFREMER	Once / day	data provider
Tide Gauge	SONEL	SHOM	3 times / week	data provider
	RONIM	SHOM	Once / hour	CDOCO (automatic only)
High frequency Radar	Vigicote - radar HF	SHOM/ACTIMAR	Once / day	ACTIMAR
Temperature and salinity networks	Island networks	IFREMER	delayed mode	CDOCO (automatic and visual)
	Recopesca	IFREMER	Once / day	CDOCO (automatic and visual)

of the swell monitoring network along the French coasts.

- Island network temperature and salinity data (Glénans, Houat, Yeu and Oléron islands in the Bay of Biscay).
- High frequency radars for wave height and direction and surface currents measurements.
- MOLIT (Mer Ouverte LIToral) fixed buoy in the bay of Vilaine (South of Brittany) for temperature, salinity, oxygen, PH, fluorescence and turbidity measurements.
- Drifting buoys deployed by IFREMER for temperature and trajectory measurements.
- Fixed meteorological buoys (Météo-France and Met-Office).

Reference data

Bathymetry determines for a large part the propagation of phenomena such as the tide or the waves. It is a key parameter for hydrodynamic modelling and it is used as embedded grids of resolution from 1 km for the definition of the boundary condition at the scale of the oceanic basin to 10 meters for coastal waters.

CDOCO is in charge of compiling and supplying the bathymetric datasets required for the hydrodynamic modelling of the coastal seas of France.

Data which have been compiled are soundings from hydrographic surveys mainly carried out by SHOM and the autonomous harbours ("Ports Autonomes") (sampled at 25m of interval) extracted from the SHOM database and bathymetric data from the French oceanographic cruises hold by Ifremer. In addition, foreign data sources such as Seazone (for UK waters) and international DB such as GEBCO have been used to cover distant areas. Various bathymetric grids have been produced in cooperation with the IFREMER Geosciences (GM) and Environmental (DYNECO) departments. At present,

CDOCO offers 3 levels of products:

- Low resolution grids (1km and 500m) for the Bay of Biscay, the Channel and the southernmost North Sea) based on simple algorithms for boundary conditions.
- A 100m grid of the French coastal waters from Spain to North Sea (depth less than 50m or distance less 40km from the coast). Regular kriging has been used to process the soundings. The computed depth is recorded with variables such as the standard deviation of the estimation of each depth or the age of source of data.
- Soundings sampled at 25m of interval of the French waters.

In 2009, the products will be extended to the Mediterranean Sea.

Satellites data

Satellites observations of sea surface temperature, ocean colour, wind and altimetry are also needed for models validation ; these data are collected and distributed by the CERSAT (Centre for Satellite Exploitation and Research, in IFREMER) data centre and do not circulate through the CDOCO.

Quality assurance

The in situ data are checked for quality either by the data provider or by CDOCO (Table 3). They are loaded in the database and then submitted to automatic/objective and visual/subjective checks according to international recommendations. These checks are performed on each profile or time series separately and also on profile grouped by sensors. As a result, the values are not modified, but a quality flag is assigned to each numerical value. In order to detect anomalies, the following automatic checks are performed on the data:

- on-land location test, for profiles or time series of sea temperature and salinity,
- global range test (measurement included between minimum and maximum values),
- spikes and gradient test to detect outliers,
- increasing pressure test for vertical profiles,
- constant values along the profile or the time series,
- pressure test by comparison with the bathymetric statistics to detect data measurement below the bottom depth,
- vertical instability when temperature and salinity are both measured.

Each profile or time–series is then checked visually, compared with pre-existing statistics, if available, and with other profiles in the same area or from the same sensor. The results of the automatic checks (quality flags) are validated or invalidated manually (Fig. 3).

Data archiving

Both observed data collected for the models and model results (hindcast, nowcast analyses and forecast) are archived.

In situ data except HF radars are loaded in the Coriolis database, whereas other data, mainly model results used or generated by PREVIMER are safeguarded on hardcopies.

A specific server has been installed for archiving: its capacity is 32 terabytes. Each day PREVIMER models produce 24 gigabytes of data that need to be archived and secured.

Data access

CDOCO is in charge of the distribution of the data. It provides personalized access to data for partners and authorized users.

Two formats have been designed for the data distribution:

- ASCII OCO: a very simple comma separated value (CSV) format, with one header line describing the columns of the file (metadata, measured parameters and their units) and then the lines of measurements.
- NetCDF OCO: based on NetCDF from Ucar, used for in situ data, straight grid, oriented grid and spectra. NetCDF OCO formats are in compliance with the CF-1 metadata convention.

For the actors of the PREVIMER operation-

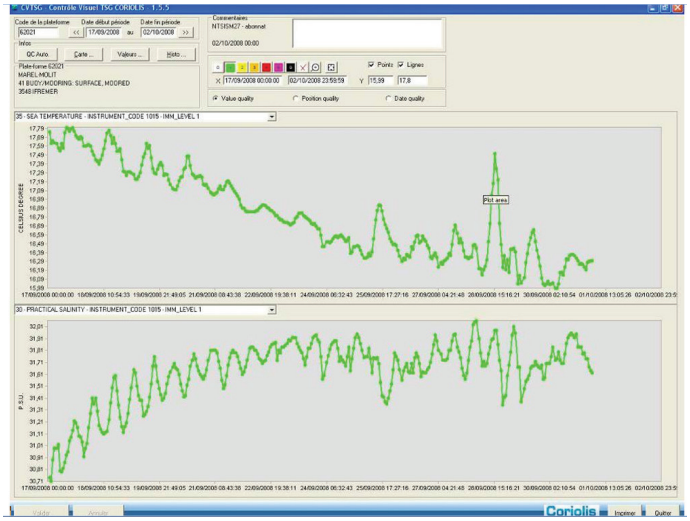

Fig. 3: manual QC of a in situ data from the MOLIT buoy (here measurements of temperature and salinity).

al system, data are distributed as follow:

- In situ data are distributed :
  - Every quarter hour at ASCII format, with one file per platform (one buoy, one ship …) containing the last 30 days of measurements. Only the new data are generated at each delivery.
  - Once a day, files of data from month M-1 and month M are distributed.
- Model results (gridded data) are available at NetCDF format to the authorized users as soon as collected by CDOCO.

The data are delivered on an ftp site, only accessible to the authorized users.

In situ data from eulerian networks are available on the Web. Public data is directly downloadable (CETMEF swell buoys, MAREL network, meteorological buoys, island network data) whereas rainfalls, river flows and tide gauges data are only available for authorized PREVIMER partners at: http://www.coriolis.eu.org/cdc/EulerianNetworks/cdcAllEulerianNetworks.asp.

Some models outputs (D1, F1) are accessible through OpenDAP thanks to a Thredds data server (TDS). They can be subset according to space-time criteria at the following address: http://www.ifremer.fr/thredds3/subcatalogs/DATA_CENTERS/OCO/OCO.html.

The results of the models are published on PREVIMER Web site (http:/www.previmer.org), where a user can find information updated daily, trend maps, forecast bulletins, archive browser, commented events.

Monitoring of the CDOCO operations

In order to provide an operational service, CDOCO has developed a web based interface which allows operators to monitor all the processes that are run in the data centre.

At present time CDOCO runs more than 40 processes which collect data, from all the different sources, including in situ data sources and model outputs.

Before running a model, it is very important to know if a data collection process has succeeded or not. In case of failure, the CDOCO operators must detect the problem as soon as possible and solve it to resubmit a new dataset to the models.

A monitoring board has been implemented to provide a synthetic view of the performance of all the data collection processes that have

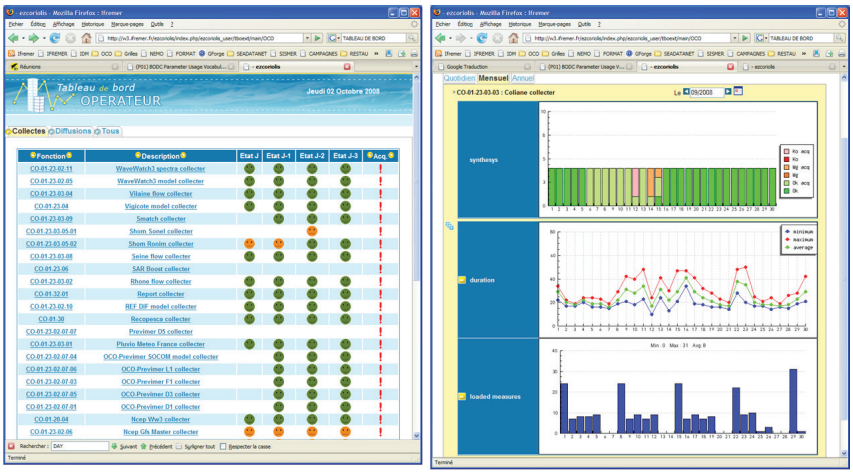

Fig. 4: Monitoring board of the data collection.

been performed during the 4 previous days.

It displays coloured faces that indicate, at a glance, whether a process has succeeded (green face), failed (red face) or if a warning (orange face) requires a special attention from the operators (Fig. 4).

Details on the error or the warning are available by clicking on the faces.

The monitoring board displays also monthly and yearly graphs on the duration of the processes of the data collection and on the number of files, stations or time series collected. The monitoring board is also dedicated to the data distribution processes, on the same principles.

Conclusion

After 3 years, the Coastal Operational Oceanography data centre (CDOCO) fulfilled its objectives with regards to data management activities (Fig. 5):

- Collecting and distributing data in near real time to the models. This includes:
  - boundary conditions,
  - forcing atmospheric and freshwater data,
  - in situ data for validation.
- Archiving data and model results

The data centre is operational but needs improvements, which will be performed during the second phase of PREVIMER, such as an extension of the services data distribution including:

- OpenDAP access to more model results, it is planned to have a better integration of these sub setting services into PREVIMER website.
- Synchronisation of all the automatic procedures and particularly communication between the data collection routines and the run of the models.
- Monitoring of the model runs and report about these runs included in the monitoring board of CDOCO.
- Define and set up the downstream services for the products distributed by the data centre.
- Catalogues of product, developed in ISO-19115 in cooperation with other major projects on operational oceanography like MerSea, MyOcean and ECOOP.
- Operate the model and the data centre based on the ITIL approach ((*Information Technology Infrastructure Library*) with definition of good practices and associated tools to produce indicators on service levels.
- Set up a centralised management of all the requests concerning PREVIMER: data re-

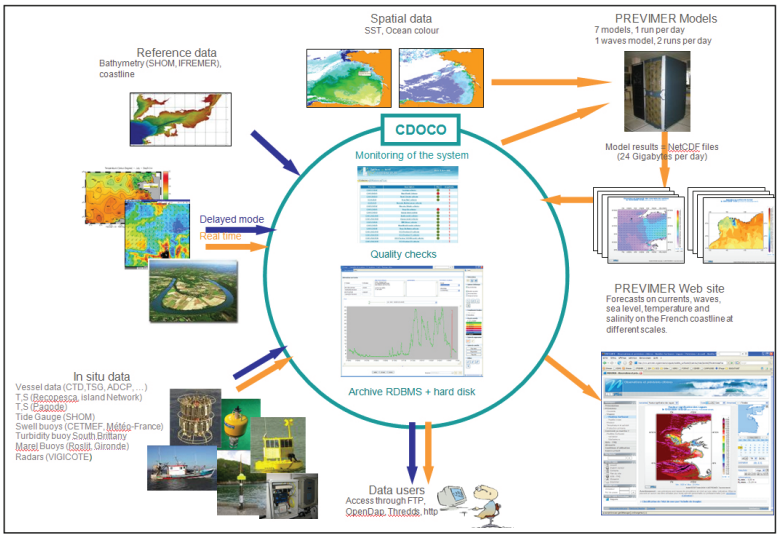

Fig. 5: Synthetic schema of the data flow in CDOCO.

quest, questions and remarks of he users, comments on the Web site, particular product needs.

Acknowledgements

The CDOCO development is part of PREVIMER program which is identified within planning contracts between the French State (represented by Ifremer and SHOM) and the Brittany region (2000-2006 and 2007-2012). The project is co-funded by the Brittany Region, the Conseil Général du Finistère, Brest Métropôle and the European Union (FEDER).

The operational developments and their implementations are conducted in the framework of a partnership agreements between public institutions (the French Naval Hydrographic and Oceanographic Service (SHOM), the French National Weather Service (Météo-France), the French public science and technology research institute (IRD), the European Institute of Marine Studies (IUEM)) and private companies (R&D departments, small to medium companies manufacturing oceanographic instrumentation)..

The data collections rely on the agreement with partners: Météo-France, SHOM, ACRI, ACTIMAR, CETMEF, CNR, DIREN, EPOC, IAV, IUEM, MERCATOR, SCHAPI.

References

CORIOLIS DATA CENTRE. Web page at http://www.coriolis.eu.org/ under Data service.

LAZURE, P. & DUMAS, F., 2008. An external–internal mode coupling for a 3D hydrodynamical model for applications at regional scale (MARS). *Advances in Water Resources 31,* p. 233-250.

LECORNU, F., DE ROECK, Y.-H. & LEGRAND, J., 2008. PREVIMER Coastal observations and forecasts. *Oceanology International (OI) 20087 conference,* London.

ODAKA, T., PREVOSTO, B.& COTTY, P., 2008. CAPARMOR CAlcul PARallèle Mutualisé pour'Oceanographie et la Recherche, Phase 1 et Phase 2. *Journée de rencontre des utilisateurs du pôle de calcul intensif pour la Mer,* Brest, Ifremer. (http://www.ifremer/fr/pcim).