OBSERVATION REQUIREMENTS FOR SCIENTIFIC ASSESSMENT OF OPERATIONAL OCEAN FORECASTING SYSTEM, AS PERFORMED IN GODAE

Fabrice Hernandez[(1)] **, Matthew J. Martin** [(2)]

[(1)] Mercator-Ocean/IRD, 8-10 rue Hermès, Parc Technologique du Canal, 31520 Ramonville St Agne, France. Email: fabrice.hernandez@mercator-ocean.fr

[(2)] United Kingdom Meteorological Office. FitzRoy Road, EX1 3PB Exeter, UK. Email: matthew.martin@metoffice.gov.uk

ABSTRACT

In the framework of GODAE, but also European funded project like MERSEA, several countries around the world have been developing an operational capacity for short term ocean dynamics prediction. Ocean forecasting systems rely on observations to provide more realistic hindcasts and forecasts, through assimilation procedures. Most of these operational groups have implemented Cal/Val procedures, based on scientific assessment of the system and products, in order to verify the realism of ocean estimates, and monitor the forecasting system in operation. Used in delayed mode, they provide a validation of numerical simulations as performed classically by the ocean modelling community. These procedures usually rely on observations, in-situ or from space. The diagnostic – metrics- can be "independent" if the set of observation has not been previously used in the assimilation. Forcing fields errors are also verified in some cases. An overview of observations used is given, as well as some requirements for future implementations of the observing system.

1. GENERAL OVERVIEW

In-situ observations have always been used by oceanographers to analyse the ocean dynamics, its variability, and verifying their theories and hypothesis. Ocean modellers conducting academic studies have also systematically relied on in-situ data, as "ground truth" to validate their numerical experiments. In both cases, in-situ data have always been too sparse to provide a continuous and complete image of the global ocean and its eddy-field variability. The era of satellite observations slightly improved the observability of ocean surface quantities, in particular for sea level (satellite altimetry), sea surface temperature (satellite radiometry), or ocean colour (satellite imagery).

In the framework of the Global Data Assimilation Experiment (GODAE, see https://www.godae.org/), operational oceanography started to emerge. Ocean forecasting systems (OFS) started to provide in real time analysis and/or forecasts. The European Union (EU) MERSEA Strand1 project (2003-2004) already intercompared [1], on a near real time basis, five existing OFS for the North Atlantic Ocean and the Mediterranean Sea. Available in-situ and satellite data were gathered, processed, and archived, in order to be used to assess the scientific quality of numerical outputs. The scientific assessment of operational products in real time was one of the core objectives of the EU MERSEA Integrated Project (2004-2008, see http://www.mersea.eu.org). Validation procedures have been designed, then implemented in the five different OFS involved: the Arctic (TOPAZ from NERSC, Norway), the Baltic (BSHcmod, from DMI, Denmark), the North East Atlantic (FOAM, from NCOF/UK-Met), the Mediterranean (MFS, from INGV, Italy), and the Global system from Mercator-Océan (France). Two 6-months targeted operational phases during the projects allowed to verify the operationally of these procedures, and more particularly assess the scientific quality of the five OFS. A particular attention was paid to use all observations available in real-time, with the goal to a) verify the quality of the operational products and b) to ensure that the different developments of the integrated project could afford the requested quality.

From these projects, an intercomparison of OFS at the international level was scheduled among international GODAE partners in 2008 [2]. It involved the majority of operational centres worldwide delivering daily ocean products, such as: BLUElink (Australia), HYCOM (USA), MOVE/MRI.COM (Japan), Mercator (France), FOAM (United Kingdom), C-NOOFS (Canada), and TOPAZ (Norway) systems. In-situ and satellite data were the only mean to certify the accuracy of one system compared to others, and it was decided to share identical observation dataset to ensure similar assessment among the different groups.

Nowadays, OFS are gradually asked to be more operational and reliable. In Europe, under the Global Monitoring for Environment and Security (GMES) program, the MyOcean project is consolidating the European OFS (see http://www.myocean.eu.org/), where assessment of the system performance and products quality is a key aspect, and where observations keep playing a major role.

What types of observations are necessary to conduct OFS scientific assessment? When are they needed? What processes need to be carried on before using observations for validation purposes? How reliable observations can be considered? How useful can be a set of observations already assimilated in the system?

All these aspects have been addressed through the projects mentioned above and some are discussed here.

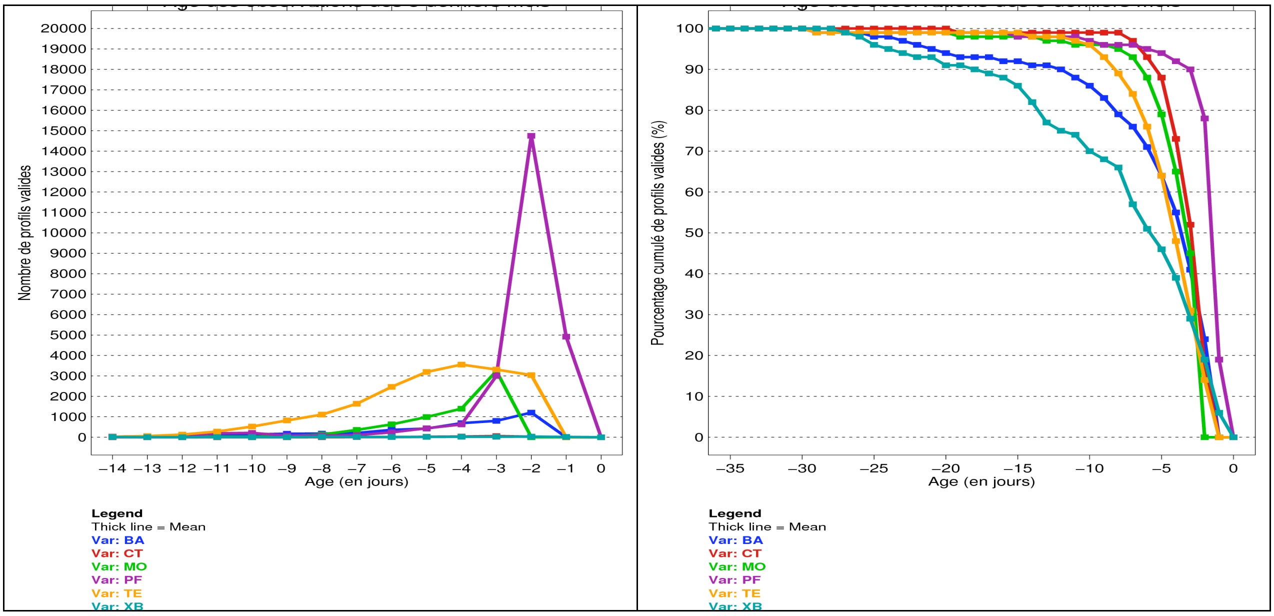

Figure 1 : Statistics of in-situ data provided by the Coriolis data centre and processed at CLS between April and June 2009. Left : age of data by type (PF, profilers / BA, Global Transmission System [GTS] BATHY messages / CT, CTDs / MO, moorings / TE, GTS TESAC messages / XB, XBTs). Right : cumulated number of data per type (courtesy of Mercator-Océan).

At present, OFS are mostly devoted to describe regional and open ocean areas, with “eddy-permitting” to “eddy-resolving” capabilities. That is, large scale O(1000km) to mesoscale O(10km) ocean dynamics. The submesoscale processes (transient filaments, fronts, waves etc…) require refined horizontal resolutions (typically 1km or less) that are rarely affordable on large areas with available computing facilities. Similarly, operational systems hardly ever represent time varying processes shorter than diurnal cycle (few hours). Usually in a given runtime they offer a description of the ocean variability over the few days neighbouring the current day. The main objective is to provide a description of the ocean dynamics in real time, with some prediction capabilities usually limited to two weeks (forecasts of atmospheric forcing are rarely provided for longer time periods). Typically, ocean hindcasts and forecasts are daily averages for open ocean, or 6-hours snapshots for regional systems. Thus, to assess their scientific quality, **observations are sought in real time to represent ocean mesoscale processes at daily scale near the surface, and slightly longer time scale below the thermocline**.

Real time availability is a specific aspect required by operational centres. Which implies that observations either in-situ or from space can be gathered in a short period of time, that necessary editing and processing procedures can also be performed “on-the-fly”, and data made available by dedicated centres to OFS. Both for assimilation and validation, **the acceptable delay for using observations is depending on the assimilation/estimations scheduling**. OFS running every day have a “collecting window” of few hours to few days before real time. For instance, the HYCOM system operated daily by the Naval Research Laboratory (USA), go back 5 days earlier. Thus it looses all data that suffer for longer delays. Weekly schedule systems like Mercator (France) goes 2 week back in time. Thus, hindcast, or best estimates are guaranteed to be evaluated with more delayed data. Delays depend on data type. For satellite observations, due to centralized and integrated structures of space centres, data availability can reach few hours. In-situ data, provided by many systems and actors suffer from a lack of availability, both for technical aspects or data policy. However, programs like Argo (autonomous profilers at depth), Surface Velocity Program (drifting buoy), tropical-mooring arrays, XBTs on ship of opportunity, were designed with a real-time transmission capability. These data are usually accessible for OFS after few hours or few days. Note that compared to the same set of data processed in delayed mode, the number of observations is usually reduced. Editing and processing performed in real time are usually eliminating a larger number of data. Figure 1 shows the monitoring performed routinely on available temperature and salinity profiles. Most profiling floats and moorings data are available in less than 5 days, but profiles transmitted through the Global Transmission System of the World Meteorological Organisation, like

TESAC and BATHY messages, arrive with one day to two weeks delay [3].

.

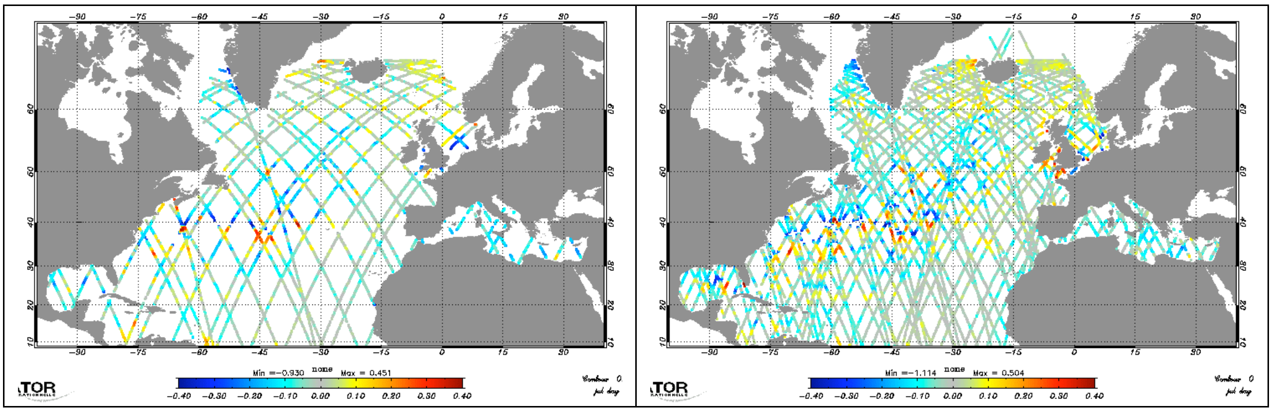

Figure 2 : Altimetric Sea Level Anomalies (SLA, heights given in meter) gathered the 24/03/2004 for the same day (left), and for the 17/03/2004, with one week delay (right). Courtesy of Mercator-Océan.

Thus, one full month can be necessary to gather all available observations provided in "near-real-time". Figure 2 gives a qualitative view of altimetric data availability in real time in 2004 over the North Atlantic Ocean, for a given day, and for the week before. Obviously, in real time, the track coverage does not allow a proper description of the eddy field.

Moreover, in real time, processing is usually less accurate than in delayed mode. For instance, satellite altimetry orbit errors need corrected models only computed after few weeks, and only delayed mode data can reach the expected accuracy of 1 cm rms. Similarly, corrections on CTDs, or moorings are obtained after laboratory analyses several weeks later. Then, correction to obtain homogenised dataset from different origins is a scientific task in itself, not performed on regular basis, and usually performed for a dedicated project like an ocean reanalysis, or ocean climatology estimation. Lot's has been learned from delayed mode processing. **Because ocean observability is poor, the lack of data can only be compensated by the combined and homogeneous use of all available data**. And it is only by gathering long time series of data than biases and errors can be identified among different type of observations. Recently, errors on XBT depth values were identified and needed to be corrected in order to reduce biases on all thermal content estimates of the ocean [4]. Note also that different type of observations of the same phenomenon can provide different aspect of the physical process, and a particular processing should be carried on to merge and use jointly different dataset. For instance, surface currents derived from satellite altimetry differ from ADCP measurements or drifting buoys velocities in many aspects: time and space scales, filtering, geostrophic vs non geostrophic components, depth of measurement. Similarly, SST from satellite radiometers or in-situ measurements capture different aspect of the ocean thermal content: "skin" vs "bulk" vs "foundation" temperature, measurement averaged over a pixel vs precise location, type of radiometer (infrared, microwave….). The GODAE High Resolution SST Pilot Project aimed to standardized SST estimations and high level mapping, but still many SST products exists and need expertise prior any validation. Alternatively one can compute the model value exactly equivalent to a given dataset (Class 4 metrics, see below). For instance, validation using drifting buoy velocity can be performed on model currents at 15 meters depth (depth of the buoy drogue), along the buoy's trajectory.

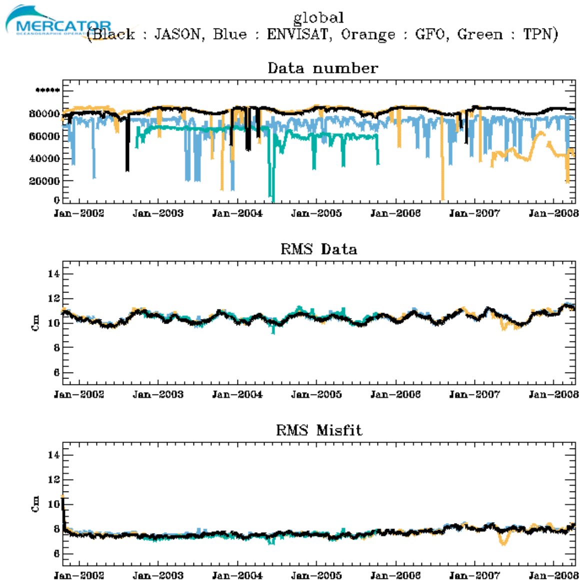

Figure 3: SLA assimilation misfits statistics from the Mercator ¼° global ocean forecasting system from 2002 to 2008. Global observation number (top) and rms (middle) are plotted from available satellite dataset (Jason-1, black / EnviSat, blue / Topex, green / GFO, orange). Bottom: rms SLA misfits. Courtesy of Mercator-Océan.

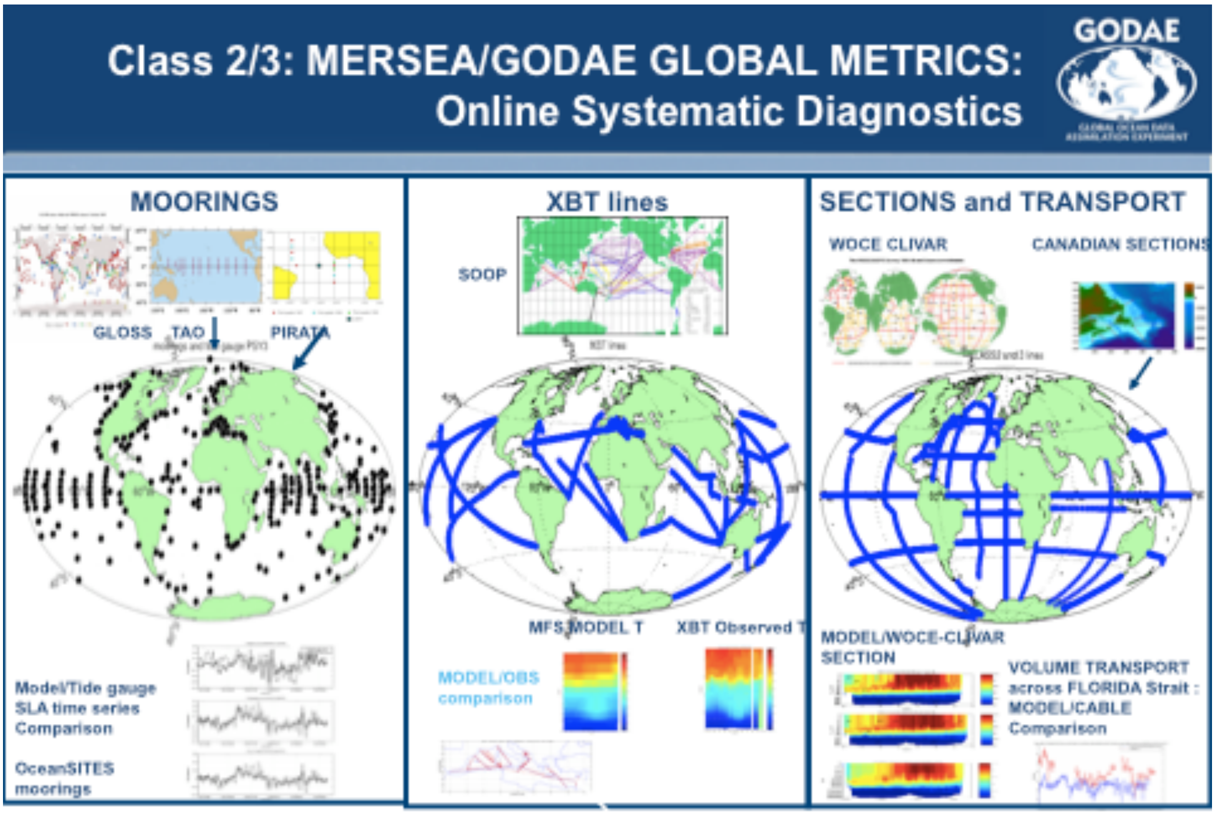

Figure 4: Summary of Class 2/3 metrics defined in the GODAE project. All available moorings, tide-gauges, XBT lines, WOCE/CLIVAR lines and others have been selected in order to define virtual sections and mooring points implemented in ocean models. Courtesy of GODAE.

For all OFS assimilating data, the first and straightforward validation tools are given by misfits (difference between the model guess or forecast and observation at the same time and location) and analysis residuals (difference between the analysis, that is, the corrected state of the ocean model after assimilation, and the same observation already assimilated). Both fields need first knowledge of observation errors in terms of measurement errors, and representativeness. Both provide a monitoring of the OFS performance and accuracy. SLA monitoring for the Mercator global system is shown in Figure 3: the “sea level truth” as given by satellite data that can be compared to Mercator outputs, taking into account the relative accuracy of each dataset.

Because all OFS do not rely on assimilation, and to take into account in a more general framework the availability of observations, and the space and time scales described by the ocean models, dedicated validation diagnostics have been designed and implemented operationally. During the MERSEA and GODAE projects, four classes of “*metrics*” taking profit of existing observations have been tested [2]. Class 1 metrics, ie 3D standardized grids of temperature, salinity, currents, mixed layer depth, sea ice quantities and fluxes, can be directly compared to climatologies, but also at the surface to satellite observations (e.g., SLA, SST, or ice concentration). By using similar Class 1 grids, several OFS can intercompare their ocean estimates with a given reference dataset. Class 2 metrics (virtual moorings and sections) are designed to match location of existing in-situ datasets as shown in Figure 4. Then each time observations are provided (e.g., an XBT sections from a merchant ship), the Class 2 diagnostic can be performed routinely, and the model variable can be compared to “ground truth”. Class 3 metrics concern derived quantities, like ocean transport, heat content, thermohaline circulation. But to get closer to data, both for hindcasts and forecasts, Class 4 metrics were designed to build up a dataset of “model values equivalent to observations” for all OFS outputs: hindcast, nowcast and forecast. Thus, forecasting skill of OFS can be objectively evaluated. Class 4 diagnostics are already implemented for temperature, salinity (observations from Coriolis data centre), sea-ice concentration (maps from sea-ice SAF), sea level (satellite altimetry from AVISO) and currents (from the Global Drifter Program). **For all these diagnostics, a particular attention is paid to use independent observations**, ie, preferably not assimilated. Ideally, instead of satellite altimetry assimilated in most OFS, tide gauge data for sea level (Class 4 metrics under implementation at Mercator-Océan), or drifter or ADCP velocities for current.

2. DISCUSSION

The definition and the use of available observations for scientific and routine assessment in operational oceanography is evolving fast. Ocean models are currently coupled with sea ice models, and sea ice diagnostics are already performed. A summary of diagnostics and observations used currently is given in Table 1. The coupling with biogeochemical models is scheduled for operations within the MyOcean timeframe (2009-2011). Ocean colour data, together with in-situ observations of biogeochemical parameters will be needed, although these data are known to be more rare than physical measurements.

Data type	Measurement
In-situ temperature	CTD (DM), XBT (RT), buoy (RT), mooring (RT/DM), TSG (DM), deep float (RT), glider (RT/DM)
In-situ salinity	CTD (DM), XCTD (DM), buoy (RT/DM), mooring (RT/DM), TSG (DM), deep float (RT), glider (RT/DM)
SST	Satellite radiometer/radar (RT), TSG (DM), buoy (RT), mooring (RT/DM)
SSS	TSG (DM), buoy (RT), mooring (RT/DM) *[SMOS, Aquarius*] (RT expected)
current	Drifters (RT), Current meter (DM), ADCP (DM) Satellite altimeter (RT), SAR (DM), HF radar (DM), derived from SST (DM), derived from deep float displacement (DM).
Sea level	Tide gauges (RT), satellite altimeter (RT), *GPS (to be tested)*
Ocean colour	Satellite imagery (RT/DM)
Sea Ice concentration, drift	Satellite (RT)

Table 1: Ocean and sea-ice physical quantities, and corresponding available observations for validation in real time (RT) or delayed mode (DM).

Moreover, regional and coastal operational systems are now emerging. Usually, a downscaling strategy is adopted: operational systems on open ocean are providing initial and boundary conditions to these models. In terms of validations two points are raised: How can be characterised routinely the impact of the large scale model errors on the regional one? And how can be addressed the validation of the regional/coastal model itself, considering the sparseness of observations? Because coastal ocean processes are characterised by shorter time and space scales of interest, impact of river run-off, physical effects of the topography and the coast lines, tidal dynamics and mixing etc… that are usually less significant on open ocean. Most of the observing system are dedicated to open ocean, for instance, satellite altimetry cannot provide accurate sea level measurements unless specific processing is performed. New technologies, like coastal radars or gliders are thus expected to provide in the near future valuable information on the shelves in order to allow the routine evaluation of operational regional/coastal systems.

3. REFERENCES

1. Johannessen, J.A., P.-Y. Le Traon, I. Robinson, K. Nittis, M.J. Bell, N. Pinardi, P. Bahurel, B. Furevik, and MERSEA Strand1 Consortium (2005), Marine EnviRonment and Security for the European Area: lessons learned from MERSEA Strand-1, in Fourth International Conference on EuroGOOS: European Operational Oceanography: Present and Future, 6-9 June 2005, Brest, France. Edited by H. Dahlin, N.C. Flemming, P. Marchand, and S.E. Petersson, in Sixth Framework Programme, 279-284, European Commission.

2. Hernandez, F., L. Bertino, G. Brassington, E. Chassignet, J. Cummings, F. Davidson, M. Drévillon, G. Garric, M. Kamachi, J.-M. Lellouche, R. Mahdon, M. J. Martin, A. Ratsimandresy, and C. Regnier (2009), Validation and intercomparison studies within GODAE, Oceanography Magazine, 22, 3, 128-143.

3. Keeley, R., Bradshaw, B. & Pazos, M. (2010). "Data Management System for Surface Drifters" in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2),* Venice, Italy, 21-25 September 2009, Hall, J., Harrison D.E. & Stammer, D., Eds., ESA Publication WPP-306.

4. Ingleby, B., and M. Huddleston (2007), Quality control of ocean temperature and salinity profiles — Historical and real-time data, J. Mar. Syst., 65, 158 – 175, doi:10.1016/j.jmarsys.2005.11.019.