



# Mercator Ocean

Ocean Forecasters



*The MyOcean Catalogue of products*

<http://www.myocean.eu/web/24-catalogue.php>

*holds 215 PRODUCTS as of July 2011, among which many Reanalysis. This catalogue will be updated on December 19 2011, with new and upgraded products as well as new download services and a viewing capacity.*

*Credits: MyOcean <http://www.myocean.eu/>*

## Editorial – October 2011 – Three of the MyOcean long time series reanalysis products

Greetings all,

This month's newsletter is devoted to three of the MyOcean long time series Reanalysis products: the In Situ temperature and salinity CORA reanalysis (1990 to 2010), the reanalysis of the North Atlantic ocean biogeochemistry (1998-2007) and the Arctic Ocean sea-ice drift reanalysis (1992-2010).

The first product described here is the In Situ temperature and salinity CORA reanalysis (1990 to 2010). A new version of the comprehensive and qualified ocean in-situ dataset (the Coriolis dataset for Re-Analysis - CORA) is released for the period 1990 to 2010. This in-situ dataset of temperature and salinity profiles, from different data types (Argo, GTS data, VOS ships, NODC historical data...) on the global scale, is meant to be used for general oceanographic research purposes, for ocean model validation, and also for initialization or assimilation of ocean models. This product is available from the MyOcean web portal (<http://www.myocean.eu/>).

The second product is the reanalysis of the North Atlantic ocean biogeochemistry (1998-2007). A system assimilating Ocean Colour SeaWiFS data during the period 1998-2007 has been designed to construct a reanalysis of the North Atlantic ocean biogeochemistry based on a coupled physical-biogeochemical model at eddy-admitting resolution. The aim of this study is, on the one hand to develop the skeleton of a pre-operational coupled physical-biogeochemical system with real-time assimilative/forecasting capability, and on the other hand to operate this prototype system for producing a biogeochemical reanalysis covering the 1998-2007 period. This product is not available from the MyOcean web portal yet.

The third reanalysis product is the 1992-2010 winter Arctic Ocean sea ice drift time series made at Ifremer/CERSAT from satellite measurements which consists of several products: the Level 3 products from single sensors and the L4 products from the combination of sensors. They are available at 3, 6 and 30 day-lag with a 62.5 km-grid size during winter. This dataset is available for oceanic and climate modelling as well as various scientific studies in the Arctic. The time series is ongoing and will continue for Arctic long term monitoring using the next MetOp/ASCAT operational scatterometers, planned to be operated for the next 10 years. This product is available from the MyOcean web portal (<http://www.myocean.eu/>).

The next January 2012 issue will be dedicated to various applications using the Mercator Ocean products.

We wish you a pleasant reading!

## CONTENT

<b>News: Workshop on Observing System Evaluation and Intercomparisons : a GODAE Ocean view/GSOP/CLIVAR event (13-17 June 2011, Santa Cruz, USA)</b>	<b>3</b>
<b>CORA3, a comprehensive and qualified ocean in-situ dataset from 1990 to 2010</b>	<b>4</b>
<i>By Cécile Cabanes, Antoine Grouazel, Victor Turpin, François Paris, Christine Coatanoan, Karina Von Schuckmann, Loïc Petit de la Villeon, Thierry Carval, Sylvie Pouliquen</i>	
<b>A multivariate reanalysis of the North Atlantic ocean biogeochemistry during 1998-2007 based on the assimilation of SeaWiFS data</b>	<b>10</b>
<i>By Clément Fontana, Pierre Brasseur, Jean-Michel Brankart</i>	
<b>Ifremer/CERSAT Arctic sea ice drift 1992-2010 time series from satellite measurements for myocean</b>	<b>20</b>
<i>By Fanny Girard-Ardhuin, Denis Croize-Fillon</i>	



## NEWS: WORKSHOP ON OBSERVING SYSTEM EVALUATION AND INTER-COMPARISONS : A GODAE OCEAN VIEW/GSOP/CLIVAR EVENT (13-17 JUNE 2011, SANTA CRUZ, USA)

Last June 2011, a workshop took place in Santa Cruz, USA, on the Observing System Evaluation and Intercomparisons. It was jointly organised by the Climate Variability Global Synthesis and Observations Panel (M. Balmaseda, CLIVAR-GSOP) as well as two of the GODAE Ocean view task teams: The Observing System Evaluation Task Team (P. Oke) and the Intercomparison and Validation Task Team (M. Martin). One of the main objectives of the workshop was to demonstrate and assess the value of observations to ocean forecast and analysis systems. Workshop report and presentations are available at the following address:

<https://www.godae-oceanview.org/outreach/meetings-workshops/godae-oceanview-gsop-clivar-workshop/>.

Specific goals of the workshop were to:

- Demonstrate the value of in-situ and satellite observations to short term, seasonal and decadal forecast systems.
- Move towards routine monitoring of the global ocean observing system (explore possibilities for “Observations Impact statements”).
- Review intercomparison of class 4 metrics from operational short-range forecast systems and seasonal prediction systems.
- Linking intercomparison and observing system evaluation monitoring activities within GODAE OceanView to those of CLIVAR.

Many studies were presented that demonstrate the value of the global ocean observing system for short term to seasonal forecasts. Overall, the main outcomes of the workshop are:

- ⇒ The ocean forecasting and reanalysis community needs a common framework for reporting their results to the observational community and agencies that are responsible for maintaining the GOOS (Global Ocean Observing System), despite the fact that the value of different type of observations (like ARGO, altimetry, Sea Surface Temperature) is already demonstrated. There was an agreement to find a way towards developing “Observing Impact Statements”.
  - ⇒ Routine monitoring of the impact of ocean observations on the forecasting and analysis systems has been developed through a range of methods, most of which are derived from data assimilation theory. A successful demonstration has been done by the UK-MetOffice. Then several operational centres state their intention to participate to the near real time OSE (Observing System Experiments) activity.
  - ⇒ The value of international intercomparison activities was highlighted by the great success of the CLIVAR-GSOP community. These intercomparisons help to clarify the common problems in the systems, and to assess the impact of the observations. It was agreed that the GODAE OceanView community should work towards the generation of a multi-model ensemble. In addition, there was an agreement to produce and intercompare various climate metrics from GOV (Godae Ocean View) and GSOP (Global Synthesis and Observations Panel).
-

# CORA3, A COMPREHENSIVE AND QUALIFIED OCEAN IN-SITU DATASET FROM 1990 TO 2010

**By Cécile Cabanes<sup>(1)</sup>, Antoine Grouazel<sup>(1)</sup>, Victor Turpin<sup>(2)</sup>, François Paris<sup>(2)</sup>, Christine Coatanoan<sup>(2)</sup>, Karina Von Schuckmann<sup>(1)</sup>, Loic Petit de la Villeon<sup>(2)</sup>, Thierry Carval<sup>(2)</sup>, Sylvie Pouliquen<sup>(1)</sup>**

<sup>1</sup> *Laboratory of Oceanography from Space (LOS), IFREMER/CNRS, Brest, France*

<sup>2</sup> *Coriolis Data Centre, IFREMER, Brest, France*

## Abstract

Coriolis is a French program which aims at contributing to the ocean in situ measurements as part of the French oceanographic operational system. It has been especially involved in gathering all global ocean in-situ observation data in real time, and developing continuous, automatic, and permanent observation networks.

A new version of the comprehensive and qualified ocean in-situ dataset, the Coriolis dataset for Re-Analysis (CORA), is produced for the period 1990 to 2010. This in-situ dataset of temperature and salinity profiles, from different data types (Argo, GTS data, VOS ships, NODC historical data...) on the global scale, is meant to be used for general oceanographic research purposes, for ocean model validation, and also for initialization or assimilation of ocean models.

To generate this new version, new and updated data have been extracted from the Coriolis database and added to the previous CORA dataset spanning the period 1990-2008. To qualify this dataset, several tests have been developed to improve in a homogeneous way the quality of the raw database and to fit the level required by the physical ocean re-analysis activities. These tests include some simple systematic tests, a test against climatology and a more elaborate statistical test involving an objective analysis method (for the validation of this dataset). Visual quality control (QC) is performed on all the suspicious temperature (T) and salinity (S) profiles and quality flags are modified in the dataset if necessary.

This Coriolis product is available on request through MyOcean Service Desk (<http://www.myocean.eu/>).

## Introduction

The Coriolis database is a real time dataset as it is updated every day as new data arrive. On the contrary, the CORA database corresponds to an extraction of all in situ temperature and salinity profiles from the Coriolis database at a given time. All the data are then re-qualified. CORA is meant to fit the needs of both re-analysis and research projects. However, dealing with the quantity of data required by re-analysis projects and the quality of data required by research projects, remains a difficult task.

Several important changes have been made since the last release of CORA2, concerning both the production procedure (to be able to release yearly reanalysis) and quality checks applied to the data. Those changes are the following:

- A new procedure is now used to produce the dataset: This procedure was set up to be able to extract only new and updated data from the Coriolis database at each new release of CORA.
- A new set of quality checks is performed on the data.
- A check of duplicates was re-run on the whole dataset
- An XBT bias correction has been applied.
- The CORA3 release has been extended for the period 1990-2010.

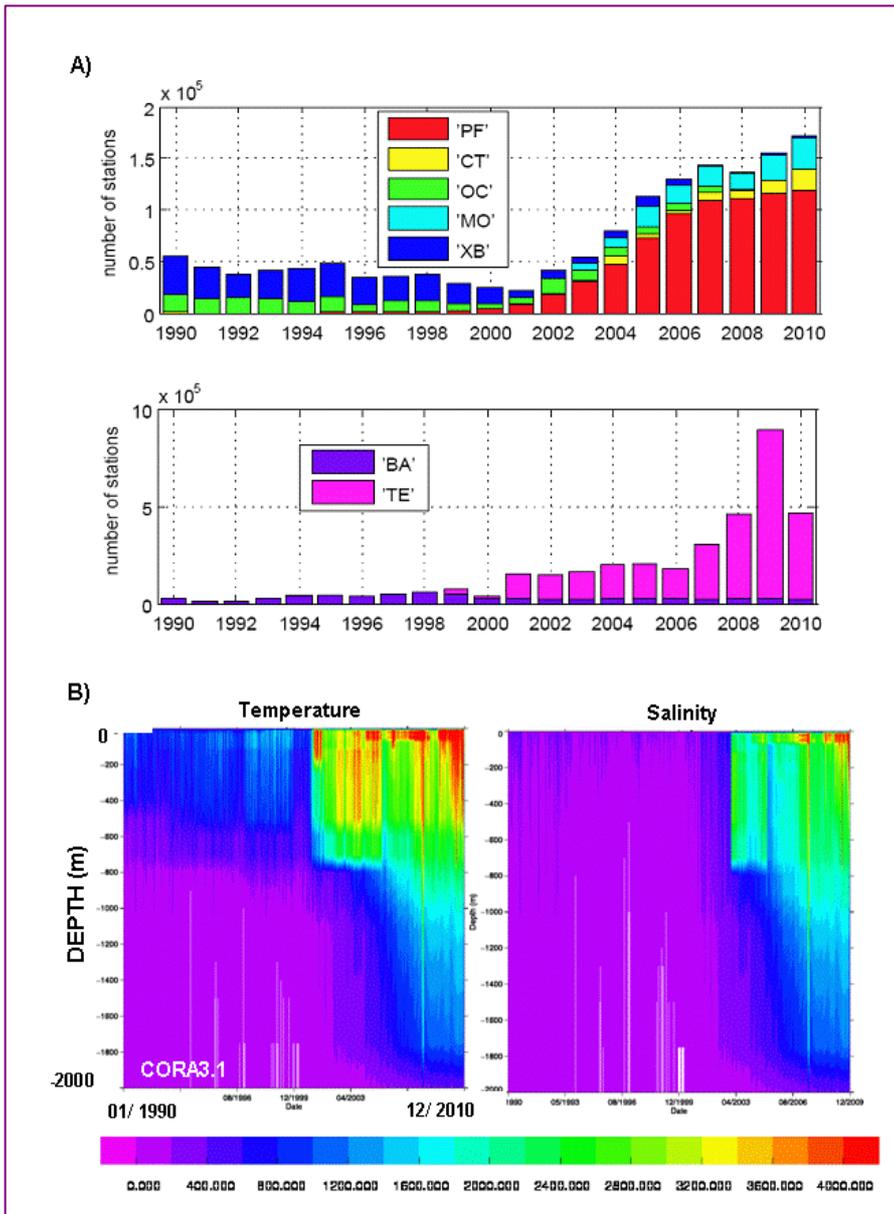


Figure 1: A) Number of profiles in the Global Ocean in the different file types in CORA3 and B) Number of temperature and salinity data per month over the global Ocean, as a function of depth (from S.Guinehut, CLS, Toulouse).

## Description of the Dataset

The Coriolis centre receives data from the Argo program, French research ships, GTS data, GTSP, GOSUD, MEDS, voluntary observing and merchants ships, moorings, and the World Ocean Database (not in real time for the last one). CORA thus contains data from different types of instruments: mainly Argo floats, XBT, CTD and XCTD, and moorings. The data are stored in 7 files types: PF, XB, CT, OC, MO, BA, and TE. The data from Argo floats directly received from DACS (PF files) have a nominal accuracy of 0.01°C and 0.01 PSU and are transmitted with full resolution. XBT or XCTD data received from research and opportunity vessels (XB files) provide an accuracy of 0.03°C to 0.1°C for temperature and 0.03 to 0.1 PSU for salinity. The CT files contain CTD data from research vessels (accuracy on the order of 0.001°C for temperature and 0.001 PSU for salinity after calibration) but also data from sea mammals equipped with CTDs (accuracy is on the order of 0.005°C for temperature and 0.02 PSU for salinity but can be lower depending of the availability of reference data for post-processing, see Boehme et al., 2009) and some Sea Gliders. Other CTD data are found in OC files and come from the *high resolution CTD* dataset of the World ocean database 2009 (Boyer et al., 2009). Mooring data (MO files) are mostly from TAO TRITON RAMA and PIRATA moorings and have accuracy comparable to Argo floats. The two last categories (TE and BA files) are for all data transmitted through the GTS (data from Argo floats not yet received at the DACS, moorings, etc...). This transmission system imposes limitations on the accuracy: data are truncated respectively two and one places beyond decimal point for TE and BA type. Figure 1 shows the number of temperature and salinity profiles in the CORA3 database for the whole period 1990-2010 and their repartition among the different file types. A large amount of data comes from the GTS as a consequence of the real time needs of the Coriolis data centre.

## Description of data processing

### Data flow

CORA data are retrieved from the Coriolis database which is constantly evolving because new data are submitted, some data are reprocessed, other data are adjusted in delayed mode (Argo data), quality flags are modified in the Coriolis database after quality checks have been performed, etc.... Basically, all new and modified data since the previous version of CORA are retrieved from the Coriolis database. This subset of new and updated data are then re-qualified. Other quality checks are performed on the whole dataset to ensure data quality consistency

### Quality checks

Data received by the Coriolis data centre from different sources are put through a set of quality control procedures (Coatanoan and Petit de la Villéon, 2005) to ensure a consistent dataset.

Beside those tests, several other quality checks have been developed or applied to produce CORA3 in order to reach the quality level required by the physical ocean re-analysis activities. These checks include some simple systematic tests; a test against climatology and a more elaborate statistical test involving an objective analysis method (see Gaillard et al., 2009 for further details). Visual quality control (QC) is performed on all suspicious temperature and salinity profiles. After these visual checks it is decided to change or not the control quality flag.

### Systematic test on new and updated data

A profile fails a systematic test when pressure is negative, T and S values are outside an acceptable range depending on depth and region, T or S are equal to zero at bottom or surface, values are constant at depth, values are outside the 10 s climatological range, if there is large salinity gradient at the surface (more than 5 PSU within 2dB) or a systematic bias. Each time a profile failed a systematic test it was visually checked.

### Tests on the whole database

A test against climatology that we call Anomaly Method was also applied. In this case, a profile failed the test if at least 50% of its data points lie outside the 5 s climatological range. This allows detecting smaller deviations compare to the 10 s check. The statistical test is based on an objective analysis run (Bretherton et al., 1976) with a three weeks window. Residuals between the raw data and the gridded field are computed by the analysis. Residuals larger than a defined value produce alerts that are then checked visually. This method combines the advantage of a collocation method since it takes into account all neighbouring sensors, and the comparison with climatology. Finally, Argo floats pointed out by the altimetric test (Guinehut et al., 2009 and <ftp://ftp.ifremer.fr/ifremer/argo/etc/argo-ast9-item13-AltimeterComparison/>) were systematically verified over all their life period and quality control flags were modified when necessary.

Quality control for the CORA3 database also includes feedbacks from N. Ferry (Mercator Ocean, Toulouse) who performed the run GLORYS2V1. This feedback provides a list of suspicious profiles detected by a comparison with the model solution. About 50% of the alerts were confirmed after visual check for the CORA3 database. Figure 2 is an example of a suspicious profile detected. This kind of feedback can help to improve our tests.

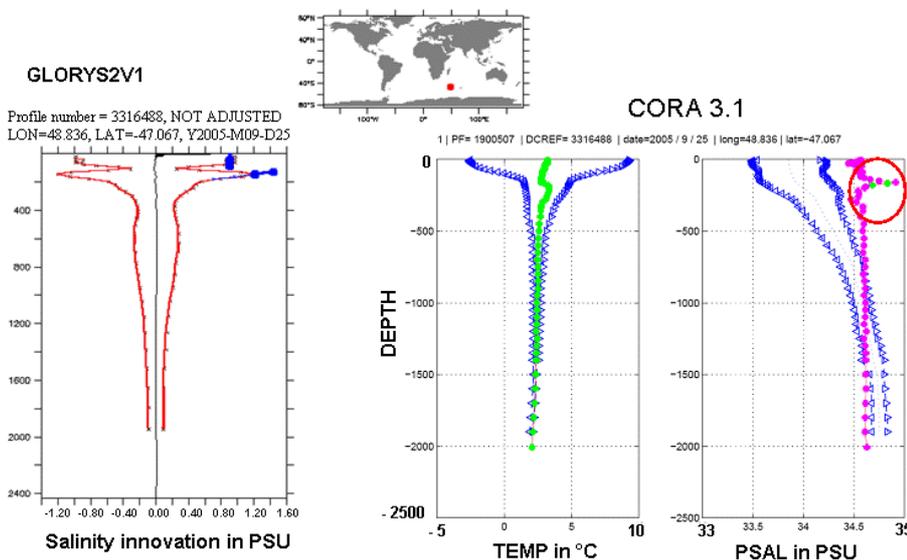


Figure 2: Example of a suspicious profile provided by the feedback of N. Ferry (Mercator Ocean, Toulouse) who performed the GLO-RY2V1 run. Salinity innovation (PSU) (left panel) and the visual check performed on this alert (right panels: temperature in °C and salinity in PSU as a function of depth).

## Elimination of redundant profiles

A last test was performed on the whole CORA dataset in order to check if double data exist. This test follows different steps:

1. All data pairs following a space-time criterion are found (the space time criterion depends on the data type).
2. For all pairs it is checked if it is a double or not: temperature and salinity profile mean and standard deviation are computed after interpolation on same depth levels, if necessary. If the differences between the profile means and standard deviations are lower than the threshold values then the pair is considered as a double.
3. If the pair is a double, then it is decided which station has to be removed from the database.
4. To make such a choice, the type of station is first considered (stations from the GTS - 'BA' or 'TE' types - are of lower precision), then the number of parameters associated to the station is considered (For example, keep the station with TEMP and PSAL instead of the station with only TEMP data) and finally the station that reach maximum depth is kept.

## XBT bias correction

Different issues with the data of expendable BathyThermograph (XBTs) exist and, if not corrected, they are known to contribute to anomalous global heat content variability (e.g. Wijffels et al. 2008). The XBT system measures the time elapsed since the probe entered the water and thus inaccuracies in the fall rate equation result in depth errors. There are also issues of temperature offset but usually with little dependence on depth.

The correction applied on CORA3 dataset is an application of the method described in Hamon et al., 2011. This correction is divided in two parts: first the computation of the thermal offset then the correction of depth. To evaluate the temperature offset and the error in depth the reference used are all the co-localised profiles (e.g. in a 3km ray, +/-15 days temporal frame, a maximum average temperature difference of 1°C and a bathymetric difference inferior to 1000m) that are not XBT and with quality flags different from 3 and 4 (suspicious and bad quality). Those references thus gather CTD, Argos profilers and mooring buoys.

### The temperature offset correction:

This correction aims to give a value of correction for each profile as a function of the year and the category of XBT: shallow XBT or deep XBT (e.g. maximal depth >=500m). The values are computed by the difference of each XBT profile with its reference profile in the layer 30-50m (below the mixed layer and where depth errors are not important enough to parasite values of temperature). Solely low temperature gradient points (e.g. <0.0025°C/m) are used to compute those corrections. XBT and reference profiles are interpolated on standard levels from 0 to 1000m and a resolution of 10m before the calculation. The final offset is the median of all those differences.

### The depth error correction:

Results of this depth correction are second order polynomial coefficients depending on the year, the depth and the category of XBT profiles. In this second step there are not two but four different categories: deep and hot, deep and cold, shallow and hot, shallow and cold (e.g. maximal depth 500m and mean temperature 11°C). To evaluate the error in depth we use the following formula:

$$dZ = \frac{T - T_{ref}}{\partial_z T_{ref}}$$

Some cursors such as a minimum number of depth errors per level and a maximum value of depth error permit to improve the quality of the median profile resulting from the raw depth errors in each of the four categories and at each level. Then we fit a second order polynomial on the median depth errors and we get three coefficients a, b and c:

$$Z_{true} - Z = aZ^2 + bZ + c$$

In the layer 30-200m, the coefficient is replaced by the mean of the depth error in order to ignore the noise due to the mixed layer.

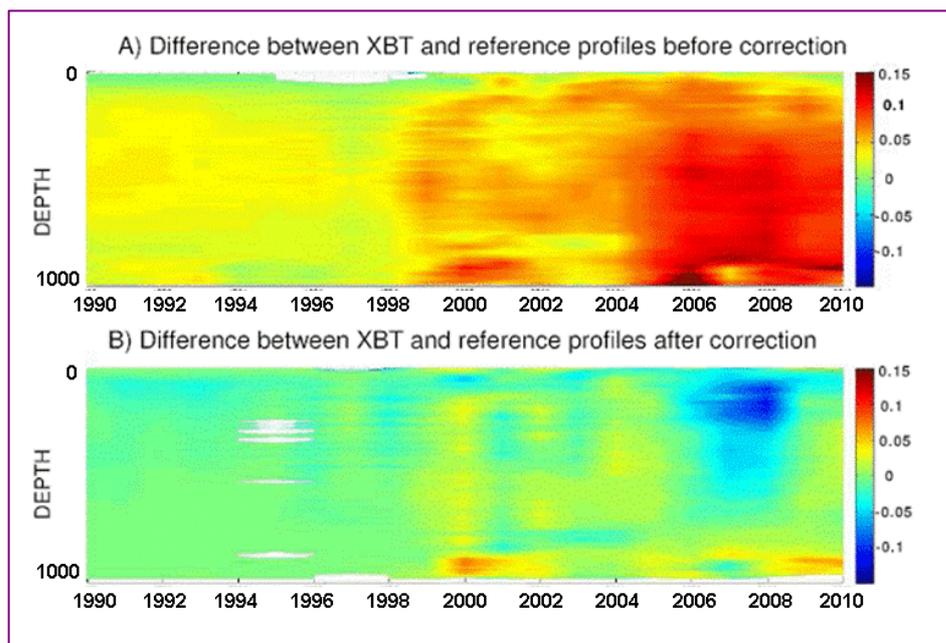


Figure 3: Temperature difference (in °C) as a function of time (years) and depth (in meters) between XBT and reference profiles before (A) and after (B) correction.

Figure 3 shows the impact of the correction on the difference between XBT and reference profiles in the CORA3 database as a function of time and depth. The large positive bias is largely reduced.

## Examples: Use of the CORA data base

### Research

Using the CORA database to estimate global ocean temperature, heat and freshwater content is a way to assess the quality of the dataset as these global quantities are very sensitive to any sensor drift or systematic instrumental bias (see Von Schuckmann and Le Traon, 2011). Although our quality controls are meant to detect such instrument problems, they can still miss small unknown drifts or bias. Comparison and sensitivity studies are thus of primary importance in the domain of climatic changes. In the Von Schuckmann and Le Traon (2011) study, Argo temperature and salinity measurements during the period 2005 to 2010 were used to estimate Global Ocean indicators (GOIs) such as global ocean heat content (GOHC), global ocean freshwater content (GOFC) and global steric sea level (GSSL). A method based on a simple box averaging scheme using a weighted mean. Uncertainties due to data processing methods and choice of climatology are estimated. This method is easy to implement and run and can be used to set up a routine monitoring of the global ocean. Over the six year time period, trends of GOHC and GSSL are  $0.55 \pm 0.1$  W/m<sup>2</sup> and  $0.69 \pm 0.14$  mm/yr, respectively. The trend of GOFC is barely significant. Results show that there is significant interannual variability at global scale, especially for GOFC. Annual mean GOIs from the today's Argo sampling can be derived with an accuracy of 0.1 cm for GSSL,  $0.21 \times 10^8$  J/m<sup>2</sup> for GOHC, and 700 km<sup>3</sup> for GOFC. Long-term trends (15 years) of GOIs based on the complete Argo sampling for the upper 1500m depth can be estimated with an accuracy of 0.03 mm/yr for GSSL, 0.02 W/m<sup>2</sup> for OHC and 20 km<sup>3</sup>/yr for GOFC - under the assumption that no systematic errors remain in the observing system.

A method that combines vertical profiles from the CORA dataset and satellite Sea Surface Height (SSH) data was also used to reconstruct the vertical structure of mesoscale eddies in the South Atlantic Ocean (see Azevedo Correia et al., 2011). This allowed estimating an annual mean eddy meridional heat flux and a volume transport from the Indian to the Atlantic Ocean and describing its interannual variability.

### Data assimilation in ocean models

An important application of such a database is also its use for ocean reanalyses. Throughout the world, several reanalyses projects are underway which aim at providing a continuous space-time description of the ocean, synthesizing the information provided by various observation types (remotely sensed and in situ) and the constraints provided by the physics of numerical ocean models. In France, global ocean reanalysis activity is a joint collaboration between Mercator Océan, Coriolis data centre and several oceanographic and atmospheric research laboratories in the framework of GLORYS (GLobal Ocean ReanalYsis and Simulations) project. This project contributes also to the production of coordinated reanalyses at the European level in the context of MyOcean EU funded FP7 project, in collaboration with Italian, English, French and Canadian partners. The goal of GLORYS is to produce a series of realistic eddy resolving global ocean reanalyses. Several reanalyses are planned, with different streams. Each stream can be produced several times with different technical and scientific choices. Stream 2 (GLORYS2) covering the period 1992-2010 has been produced using the CORA3 data set.

## References

- Azevedo Correia J.M., C. De Boyer Montegut, C. Cabanes, P. Klein : Estimation of the Agulhas ring impacts on meridional heat fluxes and transport using ARGO floats and satellite data, submitted to GRL August 2011.
- Boehme, L P. Lovell, M. Biuw, F. Roquet, J. Nicholson, S. E. Thorpe, M. P. Meredith, and M. Fedak, 2009: Technical Note: Animal-borne CTD-Satellite Relay Data Loggers for real-time oceanographic data collection, *Ocean Sci.*, 5, 685–695
- Boyer, T. P., J. I. Antonov, O. K. Baranova, H. E. Garcia, D. R. Johnson, R. A. Locarnini, A. V. Mishonov, T. D. O'Brien, D. Seidov, I. V. Smolyar, M. M. Zweng, 2009. *World Ocean Database 2009*. S. Levitus, Ed., NOAA Atlas NESDIS 66, U.S. Gov. Printing Office, Wash., D.C., 216 pp., DVDs.
- Bretherton, F., R. Davis, and C. Fandry, 1976: A technique for objective analysis and design of oceanic experiments applied to Mode-73. *Deep-Sea Res.*, 23, 559–582.
- Coatanoan C., and L. Petit de la Villéon, 2005: Coriolis Data Centre, In-situ data quality control procedures, Ifremer report, 17 pp. (<http://www.coriolis.eu.org/cdc/documents/cordo-rap-04-047-quality-control.pdf>)
- Gaillard F., E. Autret, V. Thierry, P. Galaup, C. Coatanoan, and T. Loubrieu, 2009: Quality controls of large Argo datasets, *J. Atmos. Ocean. Tech.*, 26, 337-351.
- Guinehut, S., C. Coatanoan, A.-L. Dhomp, P.-Y. Le Traon and G. Larnicol, 2009: On the Use of Satellite Altimeter Data in Argo Quality Control, *Journal of Atmospheric and Oceanic Technology*, 26, 395-402
- Hamon M., G. Reverdin, P. Y. Le Traon : Empirical correction of XBT data, submitted to JGR June 2011
- Von Schuckmann K. and P-Y. Le Traon, 2011: How well can we derive Global Ocean Indicators from Argo data?, *Ocean Sci. Discuss.*, 8, 999–1024.
- Wijffels, Susan E., J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K. Ridgway, and J. A. Church, 2008: Changing Expendable Bathythermograph Fall Rates and Their Impact on Estimates of Thermosteric Sea Level Rise. *J. Climate*, 21, 5657–5672. doi: 10.1175/2008JCLI2290.1
-

# A MULTIVARIATE REANALYSIS OF THE NORTH ATLANTIC OCEAN BIOGEO-CHEMISTRY DURING 1998-2007 BASED ON THE ASSIMILATION OF SEAWIFS DATA

By **Clément Fontana<sup>(1)</sup>**, **Pierre Brasseur<sup>(1)</sup>**, **Jean-Michel Brankart<sup>(1)</sup>**

<sup>(1)</sup> LEGI-CNRS, Grenoble, France

## Abstract

In this paper, we describe an assimilative system that has been designed to construct a reanalysis of the North Atlantic ocean biogeochemistry based on satellite ocean colour data and a coupled physical-biogeochemical model at eddy-admitting resolution. Two distinct experiments, which differ by the nature of the analysis scheme implemented in a simplified SEEK filter, have been assimilating SeaWiFS data during the period 1998-2007. A comparison of the performances obtained with the two reanalyses demonstrate the benefit of a non-linear analysis scheme that performs anamorphic transformations of the biogeochemical variables. More specifically, the non-linear analysis scheme improves the temporal description of spatially averaged surface chlorophyll concentration over the whole period of interest decreasing the RMS log error between satellite chlorophyll data and model forecasts from 1.25 to 0.74. A comparison of the reanalysed fields collocated in space and time with independent *in situ* nitrate measurements shows that the multivariate scheme allows improving objectively the nitrate concentration estimates in the upper 50 meter sub-surface layer. In that case the RMS log error is decreased from 0.74 to 0.31 in the first superficial layer of the model. Finally, a number of weaknesses are identified in the present products, providing guidance for the tuning of the next reanalysis system.

## Introduction

The observation of ocean colour from satellites is a promising source of information to better understand and monitor the ocean biogeochemistry at global scale. During the past decade, several sensors (e.g. SeaWiFS, MODIS, MERIS) have been deployed to measure globally and daily the ocean water leaving irradiance at different wave lengths in the visible spectrum. Empirical optical algorithms calibrated according to the sensor's properties enable the estimation of biogeochemical quantities such as chlorophyll or chromophoric dissolved organic matter (CDOM) concentrations.

In the framework of the Global Ocean Data Assimilation Experiment (GODAE) program and the EU MyOcean project, a number of research initiatives are being conducted to integrate biogeochemical modelling components into state-of-the-art operational global ocean models (Brasseur *et al.*, 2009). However, since the processes governing the biogeochemistry of the ocean are still poorly known, the coupled physical-biogeochemical models (CPBM) in place today still exhibit large errors when compared to available field data. Therefore, the optimal combination of ocean colour observations with biogeochemical models through data assimilation is required to realistically reconstruct the variability of marine productivity in space and time. Several difficulties must be faced to tackle this challenge, mostly related to the inaccuracy of the data set, the model non-linearities and the computational resources needed to run coupled experiments. The aim of this study is, on the one hand, to develop the skeleton of a pre-operational CPBM system with real-time assimilative/forecasting capability, and on the other hand, to operate this prototype system for producing a biogeochemical reanalysis covering the 1998-2007 period. Such a reanalysis could be used to initialize large-scale CPBMs as well as to provide realistic boundary conditions for higher-resolution environmental modelling studies at regional scales. Further, the system could also be used for Observing System Simulation Experiments dedicated to the optimization of observation/sampling strategies (see e.g. Schiller *et al.*, 2004). So far, the majority of scientific publications dealing with large-scale ocean colour data assimilation have been focused on the description of idealized twin experiments (e.g., Carmillet *et al.*, 2001; Doron *et al.*, 2011), short-term multivariate experiments (Natvik and Evensen, 2003; Ourmieres *et al.*, 2009) or long-term univariate experiments (Nerger and Gregg, 2007, 2008). We propose here to perform a realistic, multivariate data assimilation experiment covering several years, in order to investigate the strength and limitations of a sequential, reduced-order statistical updating scheme to control the biogeochemical components of a 1/4° CPBM of the North Atlantic. To our knowledge the eddy-admitting resolution of the model used here makes our investigation unique and fully consistent with the physical models of the Mercator operational suite. This paper describes first the technical set up of the assimilation experiment variants tested here, focusing on the model features and the key ingredients of the assimilation system. The impact of the assimilation process on the biogeochemical state variables modeling is then discussed along the simulation period. Finally, we conclude with a number of perspectives to be explored after this first investigation.

## Numerical tools

In this section are described the physical and biogeochemical components of the CPBM configuration, the linear and non-linear updating schemes implemented in the assimilation setup and the main features and associated errors of the assimilated data.

### The physical model

The North Atlantic ocean circulation configuration is based on the NEMO code (Barnier *et al.*, 2006), which is a primitive equation model that uses the free surface formulation. The prognostic variables are the three-dimensional velocity fields, the thermohaline variables and the sea surface elevation. Vertical mixing of momentum and tracers is modelled according to the TKE turbulence closure scheme, while convection is parameterized with enhanced diffusivity and viscosity. The model domain covers the North Atlantic basin from 20°S to 80°N and from 98°W to 23°E (Figure 1). Buffer zones are defined at the southern, northern and eastern (Mediterranean) boundaries. The horizontal resolution is  $\frac{1}{4}^\circ$ , which is considered as eddy-permitting in the mid-latitudes. The vertical discretization is done on 45 geopotential levels, with a grid spacing increasing from 6 m at the surface to 250 m at the bottom. The model is forced by ERA-INTERIM atmospheric data from the ECMWF reanalysis, using bulk formulations. The atmospheric 10-meter humidity, temperature, and wind fields are updated every 6 hours. Short- and long-wave radiations are updated every days while precipitations are updated every months.

### The biogeochemical model

The biogeochemical component of the CPBM is LOBSTER (LOCEAN Biogeochemical Simulation Tool for Ecosystem and Resources), a nitrogen-based model that contains six prognostic variables: nitrate, ammonium, phytoplankton, zooplankton, detritus and semi-labile dissolved organic matter. The chlorophyll concentration is considered as a diagnostic variable, estimated using the phytoplankton concentration through a space and time-dependant Chl/N ratio (Levy *et al.*, 2005). The LOBSTER model updates the biogeochemical concentrations at the same frequency as the baroclinic physical timestep (*i.e.* 40 minutes). Biogeochemical variables are then advected and diffused following the oceanic circulation computed by the physical model. A more detailed description of the CPBM can be found in Ourmières *et al.* (2009).

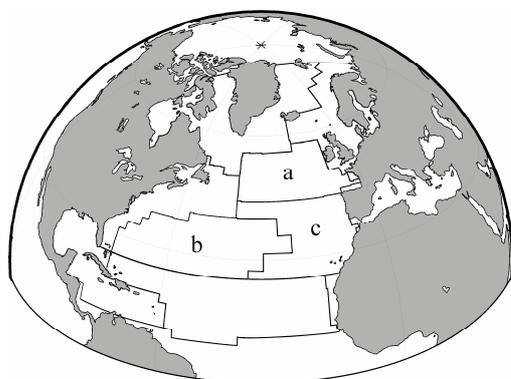


Figure 1: Model domain and biogeochemical regions a, b and c adapted from the Longhurst (1995) classification.

### The assimilation approach for Ocean Colour Data

As every ocean simulation, the numerical system used for this study is subject to a large variety of modelling errors of different nature. The main error sources which are suspected to affect both the physical and biogeochemical model response including their mutual coupling, are associated to imperfect initial conditions and atmospheric fields (themselves derived from a numerical model), parameterizations of unresolved physical or biological processes (which, in general, involve uncertain parameters or process simplifications), numerical discretization errors, etc. In the present implementation however, it is assumed that the ocean colour observations available to improve model estimates only contain the footprint of errors on the biogeochemistry. As a result, the physical model state is left unchanged by the assimilation of ocean colour data, and the ocean circulation fields are therefore treated as being assumed perfect.

The assimilation method is implemented using a sequential algorithm derived from the Singular Evolutive Extended Kalman (SEEK) filter (Pham *et al.*, 1998): the biogeochemical model trajectory is corrected intermittently by computing statistical updates of the state vector using a combination of available observations and model predictions weighted according to their respective uncertainties. The SESAM tool (Brankart *et al.*, 2011) is used to perform the algebraic operations related to the assimilation scheme, such as the computation of Empirical Orthogonal Functions (EOFs), innovation vector or analysis step. The choice of the fixed variant of the SEEK filter (Brasseur and Verron, 2006) is guided by its relative low cost in terms of numerical resources, an advantage that allows the performance of much longer experiments compared to stochastic ensemble-based

methods that require the explicit propagation of estimation errors through the model dynamics (e.g. EnKF). Nevertheless, as ensemble methods represent promising perspectives in the set-up of operational systems, our assimilation strategy has been carefully chosen to easily allow upgrades toward ensemble based methods.

### The assimilated data set

The data assimilated in this study are chlorophyll concentrations derived from the SeaWiFS sensor using the OC4 algorithm, applied to remotely-sensed water leaving irradiance measurements (Feldman and McCain, 2010). In order to facilitate the assimilation process, satellite maps were systematically remapped onto the model grid prior to assimilation. We choose to assimilate the SeaWiFS data set because this sensor is the one offering the longer time-series of ocean colour data so far (9/1997-12/2010) while others equivalent spatial sensors (i.e. MODIS and MERIS) were launched later (2002).

The part of water-leaving irradiance measured on top-of-atmosphere by satellite sensors such as SeaWiFS is typically only 5-20% of the total signal (Lavender *et al.*, 2005). The rest of the signal is due to interactions between light and earth atmosphere (e.g. Rayleigh reflectance, aerosol reflectance). Therefore in order to obtain accurate satellite data in the visible spectrum, one needs to apply atmospheric corrections to estimate the water-leaving irradiance. These corrections are a main source of errors when retrieving physical and biological quantities from the original water leaving irradiance. More specifically, the OC4 algorithm used here contains empirical parameterizations which can yield erroneous estimations of the surface ocean chlorophyll concentration, e.g. due to the assemblage of different phytoplankton types with different optical properties. Considering this, it is admitted that the root-mean-square log error of SeaWiFS chlorophyll estimation is, on average, 30% of the measured signal in deep sea waters (Gregg and Casey, 2004).

## Protocol of the assimilation experiments

### Initialization of the coupled model, simulation period and assimilation setup

The coupled model is initialized on January 1<sup>st</sup>, 1998 after a prior spin-up of the circulation model lasting 16 years. The biogeochemical model is then spun up on a 2 years period starting from the Levitus climatology for nitrate (Garcia *et al.*, 2006) and spatially homogenous values for phytoplankton, zooplankton, dissolved organic nitrogen, detritus and ammonium (see Ourmières *et al.*, 2009). The coupled model is then run without data assimilation of ocean colour data during a 9 years period until January 1<sup>st</sup>, 2007. This simulation will be referred hereafter as the “*free*” run.

In addition to the free run, two simulations using the data assimilation system are performed. Each of these runs assimilates satellite chlorophyll maps that are temporally binned every 8 days preceding the analysis date. The state vector used in the assimilation scheme is composed of every prognostic biogeochemical variable of the three-dimensional grid (phytoplankton, zooplankton, nitrate, ammonium, dissolved organic nitrogen, detritus). Thus we perform here a multivariate analysis where not only the observed variable (phytoplankton) is corrected but every components of the biogeochemical model are impacted. Physical variables of the model are not included in the assimilation process and thus are not impacted.

For the first run, the analysis update is performed using the conventional SEEK filter analysis scheme applied to the whole biogeochemical state vector. This simulation will be referred hereafter as the “*linear*” run. For the second run, an anamorphosis transformation is applied to the biogeochemical state vector, the observation vector and the error covariance matrix. The general idea is to transform the marginal probability density functions (pdfs) of the state variables into pdfs that are close to Gaussian. This is achieved by performing a change of variables (anamorphosis) separately for each single variable of the state vector. The anamorphosis transformation we use here is fully detailed in Béal *et al.*, (2010), while the effect of this transformation on the representation of model uncertainties is further discussed in Brankart *et al.* (2011). This simulation will be referred hereafter as the “*anamorphosis*” run.

The background error covariance matrix of the analysis scheme is represented using an EOFs basis obtained from the free run variability. For each analysis date, a new set of EOFs is considered from an ensemble of state vectors spread in time around the analysis date with a 2 days interval. This ensemble is constructed as follow: for a given day of the year, the model state vectors on a 2 months period surrounding the assimilation date are retained, and that for years 1999 to 2005 (in order not to exceed the free simulation period boundaries). Each of these temporal ensembles contains 210 members that are used to compute the error covariance matrix by keeping the 20 dominant EOFs. Those EOFs are the leading directions retained to perform the state vector update. It is noticeable here that the anamorphosis transformation is applied using a temporal ensemble rather than a stochastic ensemble (see below), which has never been tested yet, to our knowledge.

SeaWiFS chlorophyll concentrations are converted into phytoplankton concentration using the Chl/N ratio computed by the coupled model. These phytoplankton maps are then assimilated in the coupled model and considered as direct measurements of the phytoplankton concentration in the upper surface layer. The error associated to every single piece of data is set to 30% of the measured signal, in agreement with the admitted SeaWiFS averaged error in deep sea waters. The spatial coverage of the data is a common issue for assimilation experiments, raising the question of

the spatial correlations of the signal. As an example, many extreme latitude regions are not visible by satellite ocean colour sensors during several month of the year (winter in the northern hemisphere). In those conditions, a local analysis scheme is required to avoid the effect of spurious correlations between mid-latitude and high latitude regions (even though possible connections between distant regions may exist in reality). Considering this, the choice was made to implement a local analysis scheme with a short influence radius for every distinct data. The horizontal e-folding radius of influence has been set to 2 grid points and the cut off radius to 5 grid points. This value was chosen as it is compliant with the order of magnitude of meso-scale features in the mid-latitude regions.

## Results and system assessment

### Chlorophyll spatial distributions

Figure 2 shows a sequence of surface chlorophyll maps temporally binned on a 60 days period during the year 2006 (first row: day 1 to 60; second row: day 61 to 120; etc.). These maps are shown for: (a) the free run, (b) the SeaWiFS data, (c) the linear run and (d) the anamorphosis run. The free run shows a number of systematic differences compared to the SeaWiFS data. The chlorophyll bloom starts slightly too late in the free run (second row), while the signature of an horizontal structure appears in the Gulf Stream area that is not present in the data (third row). Consistently with this observation, the available nutrients are quickly consumed (not shown) resulting in a strong increase of phytoplankton concentrations. During the peak of the free run bloom, the chlorophyll concentrations seem to be overestimated at high latitudes. When the nutrients are fully consumed, the chlorophyll concentration decreases quickly (fifth and sixth rows) while the SeaWiFS data exhibit more significant values. To summarize, the bloom represented by the free run starts too late when compared to satellite data, while nutrients are consumed too quickly. It is important to notice here that in spite of these differences, the annual chlorophyll evolution is relatively well described (spring bloom, subtropical gyre). This is an important element to be noticed, as the free run variability provides the basic input to the EOFs computation required by the assimilation scheme.

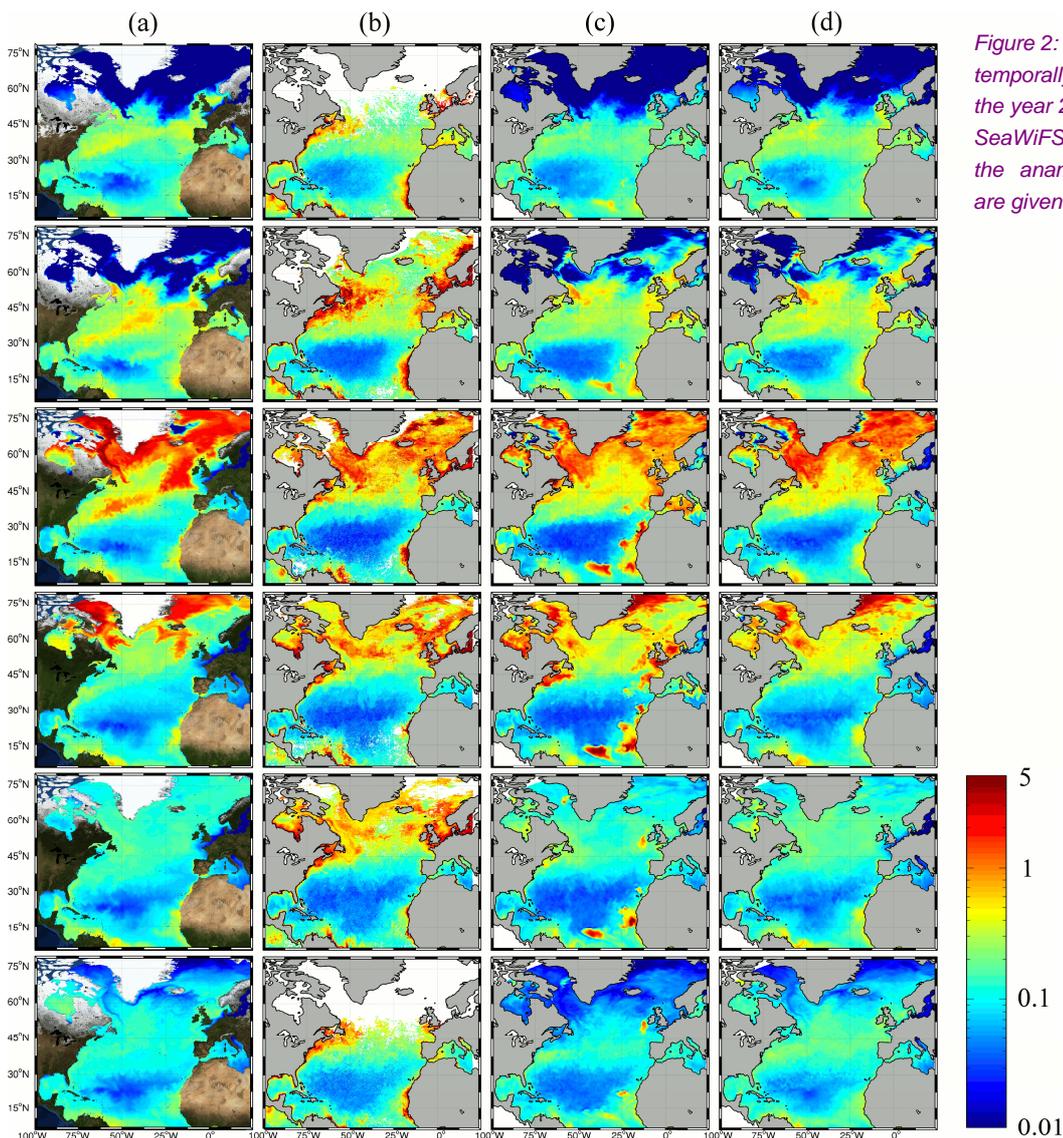


Figure 2: Surface chlorophyll distributions temporally binned on 60 days period for the year 2006 for: (a) the free run, (b) the SeaWiFS data, (c) the linear run and (d) the anamorphosis run. Concentrations are given in  $\text{mg(Chl).m}^{-3}$ .

When looking at the assimilation results, the bloom starts at a more realistic date, and the horizontal pattern discussed here above (Gulf Stream) is not present anymore (c and d). During the bloom, the modelled values are in agreement with the data, both in the linear and the anamorphosis experiments (third and fourth row). However, the chlorophyll values are still underestimated by the end of the year, like in the free run.

The error covariance matrix is computed with free run model outputs on a 2 months period running windows over years. Thus if the free run variability is much lower than the data variability on a 2 months period, the EOFs set does not contain the intrinsic dynamics allowing to satisfyingly correct the instantaneous state. This is actually what happen by the end of the year (fifth and sixth rows) when comparing the free run (a) to satellite data (b). This issue could be overcome by increasing the temporal windows on which the EOFs are computed. This will allow to catch more variability in the EOFs set used for the SEEK analysis. Nevertheless ones must be aware that increasing this temporal window could also lead to erroneous corrections as it will include in the correction base past or future correlations, not present in the instantaneous system state. From a stochastic point of view, it will include in the correction base numerous states having an extremely low probability to happen (on the edges of the temporal window).

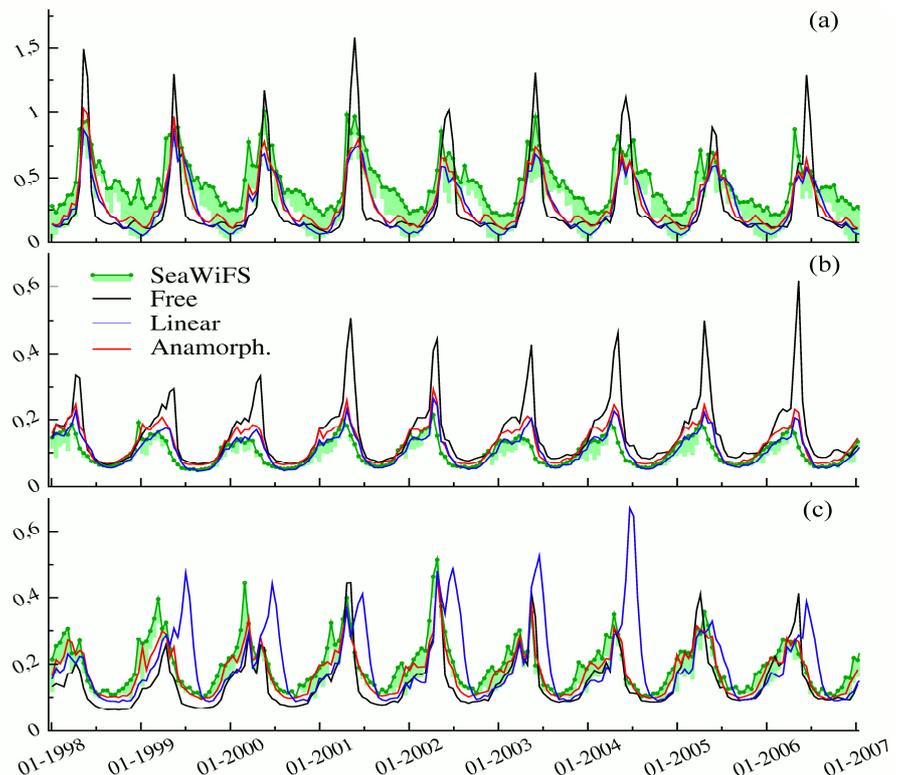
The impact of the assimilation scheme is weaker in coastal waters; this is a somehow expected result as the NEMO model is not fine-tuned to describe precisely small scale oceanic circulation features in the coastal zone. In addition, river plumes are simply considered as runoffs with no nutrient input associated to the river discharges. These processes, however, can play a major biogeochemical role in the coastal area, as demonstrated in the context of ocean colour data assimilation experiments by Fontana *et al.* (2009, 2010).

Some spots of high chlorophyll concentration appear in the linear run (c; 0°-15°N; fourth row) while these patterns are not present either in the other run, nor in the data. This is a likely consequence of the multivariate analysis that is performed here. In highly productive zones, such as in front of the Senegal Rivers (15°N, 25°W) where upwelling and nitrate inputs govern the biogeochemical response, it may happen that erroneous corrections applied to non-observed variables (e.g. nutrients) result into spurious phytoplankton growths. These patterns do not appear in the anamorphosis run (d), attesting that the corrections applied to non-observed variable are more realistic when using a non-linear assimilation scheme.

### Chlorophyll temporal evolution

Figure 3 shows the temporal evolution of the horizontally averaged chlorophyll concentration in the first layer of the coupled model over the simulation period. Both data and model outputs were temporally binned into 16-day periods. The temporal evolution is shown for 3 open-sea biogeochemical regions (a, b and c; see Figure 1) adapted from the Longhurst (1995) classification. The green fringe under the SeaWiFS curve is an indicator of the satellite data spatial coverage over the considered region. The fringe thickness is zero when the whole considered region is covered by the satellite data. Then the fringe thickness increases linearly with respect to the number of missing observations to reach 100% of the considered data value when almost no data are available. Of course when no data are present at all, no dots and no fringe are drawn.

Figure 3: Temporal evolution of the horizontally averaged chlorophyll concentration in the first layer of the coupled model for the different biogeochemical regions (see Figure 1). The chlorophyll concentration is expressed in  $mg(Chl).m^{-3}$ .



It appears on Figure 3 that for high-latitude regions (a) the bloom as modelled by the free run starts too late, does not last as long as it should when compared to data, and values are overestimated at the peak of the bloom, as discussed in the previous paragraph. One can see that, during the bloom period, the spatial SeaWiFS coverage is high (small green fringe) allowing a well-controlled chlorophyll concentration for both linear and anamorphosis assimilation runs. It is unfortunately not the case in winter when almost no data are present during several months of the year for high latitude regions. Thus in winter period and for high latitudes, it appears that not enough data are integrated in the local assimilation process to allow a significant improvement of the mean surface chlorophyll. For the mid-latitude region b, the bloom period is well reproduced by the free run, while the data are largely overestimated during the peak of the bloom, mainly due to the spurious Gulf Stream pattern discussed before. It is shown that, for this region, the linear and anamorphosis runs are able to efficiently constrain the chlorophyll evolution. Further, a drift of the chlorophyll concentration appears after several years of free simulation while the signal is more stable when using the assimilation process. For the mid-latitude region c, the free run performs well both in terms of timing and maximum value of the bloom. The anamorphosis run outstandingly improves the chlorophyll description all along the simulation period. The results are totally different when considering the linear run. Indeed, after the first year of simulation, the chlorophyll significantly diverges from the data and the others simulations. This could be the consequence of the high chlorophyll spots appearing in that region as described in the previous paragraph. The unrealistic corrections applied on non-observed variables in the case of the linear run are responsible for side effects on the biogeochemical dynamics that make the method fail for pluri-annual simulations. This demonstrates how the anamorphosis transformation can be a critical point in setting up an efficient and stable assimilation system for CPBMs using satellite ocean colour data.

### Surface chlorophyll concentrations forecast

During the assimilation period, every distinct assimilated observation was compared to its 8-days forecast equivalent for each modeled situation. This comparison is based on 410 SeaWiFS images and more than  $1.5 \cdot 10^7$  distinct data. Figure 4 shows the probability density function (pdf) of the  $\log(C_{\text{SeaWiFS}}/C_{\text{model}})$  function where  $C_{\text{SeaWiFS}}$  and  $C_{\text{model}}$  stand for the SeaWiFS and the model concentration (resp.). To compute this pdf, only open sea data entered the computation. Open sea regions were defined as area with a bathymetry deeper than 500 meters.

It appears on Fig. 4 that the chlorophyll content forecasting is strongly biased by the free run, with an overestimation of the measured values by the model (negative values) implying a RMS of 1.25. This general behavior is mainly due to the overestimation of the maximum chlorophyll concentrations in high latitudes regions during the spring bloom, as discussed before. The linear run shows a significant improvement of this comparison by centering the pdf closer to 0 while a light trend to overestimation is still visible. In the case of the linear run the computed RMS was 0.86. The result is even better when considering the non-linear run. Indeed, while the non-linear pdf is unchanged for extreme values (log error lower than -2 and higher than 2) when compared to the linear one, the pdf is better centered on 0 for low values of log error. Concerning the non-linear RMS, result was slightly better than the linear run with a value of 0.74 and thus a reduction of 41% compared to the free run RMS.

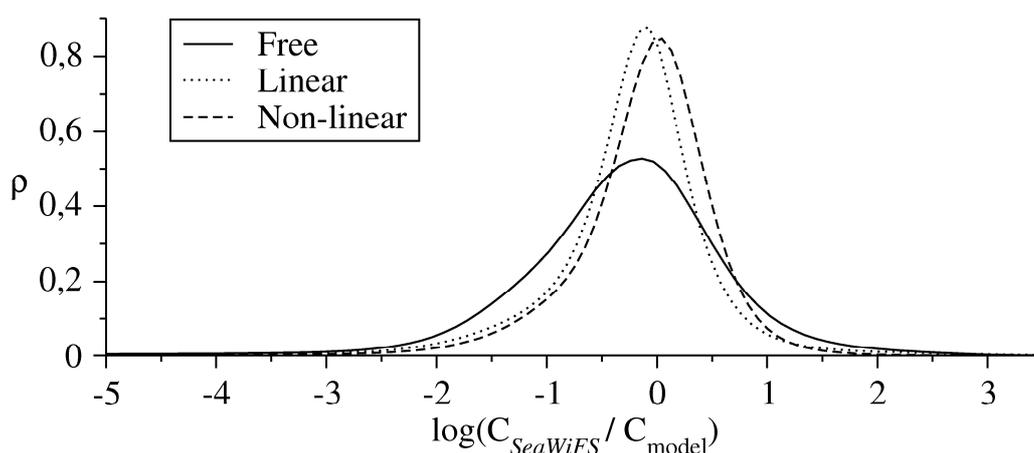


Figure 4: Probability density function of the  $\log(C_{\text{SeaWiFS}}/C_{\text{model}})$  function computed for every distinct forecasted satellite data.

### Comparison of nitrate concentration with an independent data set

The *World Ocean Atlas 2009* (WOA09) data set has been used to assess and validate the assimilation experiments with respect to independent, unassimilated *in situ* measurements of nitrate. The historical data reported in this atlas were obtained from the National Oceanographic Data Center (NODC) and the World Data Center (WDC) archives that include the measurements gathered as part of the GODAR and WOD project (Boyer *et al.*, 2006). The horizontal, vertical and temporal coverage of the available data set allows the comparison with the outputs of the 3 simulations (free, linear, anamorphosis) collocated in space and time.

The data were grouped on running vertical intervals (0-5; 5-10; 10-50; 30-70; 50-90; 70-110; 90-130; 110-150; 130-170, in meters). On these data groups, the logarithmic Root Mean Square error was computed as follows:

$$\text{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\log_{10}(C_{\text{mod}}) - \log_{10}(C_{\text{in situ}}))^2}$$

where  $C_{\text{mod}}$  and  $C_{\text{in situ}}$  stand for the modelled concentration and the measured concentration (resp.) and  $N$  is the number of data on which the computation is done. For each depth interval listed above  $N$  is equal to (resp.): 357, 286, 1100, 886, 745, 1544, 1554, 487, 62.

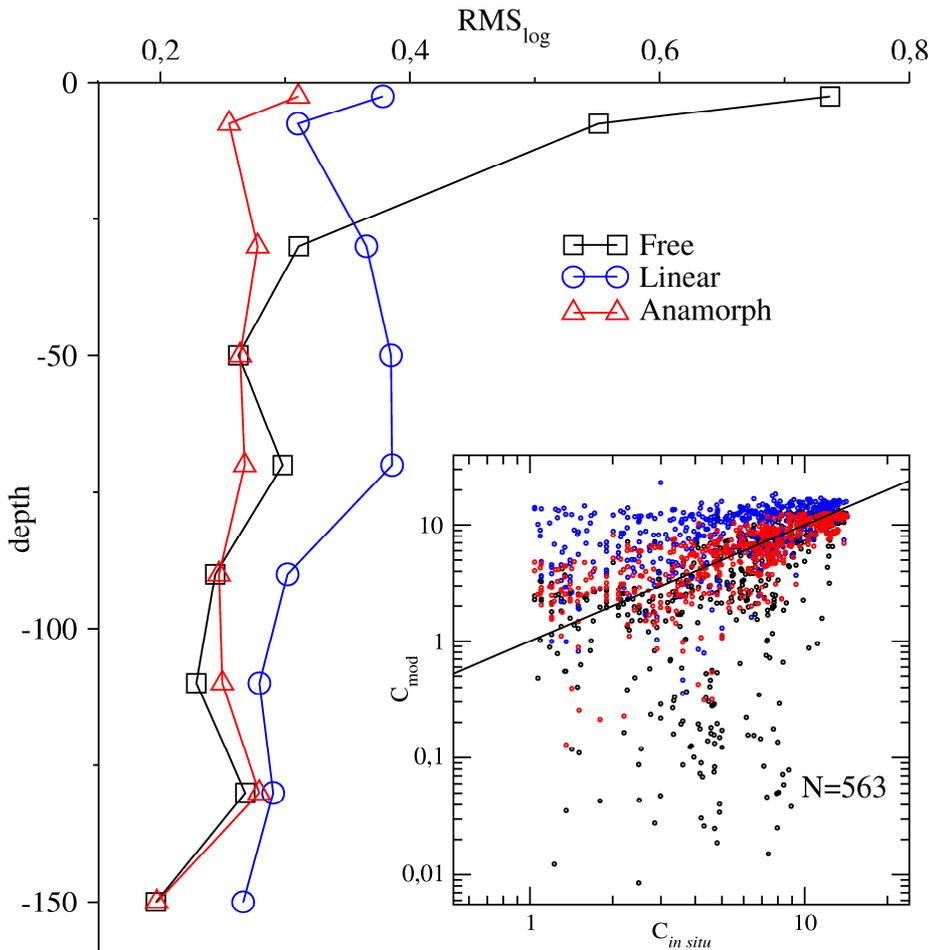


Figure 5: Nitrate  $\text{RMS}_{\log}$  misfit in the free, linear and anamorphosis runs, as a function of depth (in meters) estimated using all available in situ measurements higher than  $1 \text{ mmol.m}^{-3}$  contained in the WOA09 data set. Frame: model versus in situ scatter plot for the depth interval 0-10m. Concentrations are given in  $\text{mmol}(\text{NO}_3).\text{m}^{-3}$ .

Figure 5 shows the vertically distributed RMS for all available nitrate data higher than  $1 \text{ mmol.m}^{-3}$  contained in the WOA09. The RMS curves are shown for the three situations: free, linear, and anamorphosis. It appears that the assimilation of satellite chlorophyll data reduces the error in the upper layers of the water column, up to the 10-50m interval. While both the linear and anamorphosis runs improve the accuracy of the nitrate estimates in the upper part of the water column, the anamorphosis performs much better than the linear one. In the anamorphosis run, the RMS is decreased from  $0.74$  to  $0.31 \text{ mmol.m}^{-3}$  in the 0-5m interval. The scatter plot shown in a frame represents the model versus *in situ* values in upper 10m layer (same colours as the RMS plot). It is shown that a number of nitrate values, largely underestimated in the free run (black dots in the frame of Fig. 5), are satisfyingly corrected in the linear run. On the other hand, large overestimations appear in the linear run case, especially for values close to  $1 \text{ mmol.m}^{-3}$  (blue dots in the frame of Fig. 5). This overestimation of nitrate is a possible explanation for secondary effects discussed above on chlorophyll dynamics at mid-latitudes regions. The anamorphosis run corrects underestimations in the same way that the linear run but overestimations of values close to  $1 \text{ mmol.m}^{-3}$  disappear (red dots in the frame of Fig. 5), explaining the good performance of the method on the 0-5 and 5-10 m depth interval with respect to the RMS. On the deeper part of the euphotic layer (from about 50 to 150 m.), the linear run clearly increases the RMS when compared to others situations. Efficiency of the anamorphosis run is mixed although differences with the free run remain nevertheless weak. The RMS profiles shown on Figure 5 attest that, in the context of this study, the assimilation of sea surface chlorophyll concentration is not sufficient to control the open ocean biogeochemistry in the water column below some tens of meters depth.

As already stated, the comparison discussed above was made with available nitrate data for measured concentrations higher than  $1 \text{ mmol.m}^{-3}$ . Concerning the data lower than  $1 \text{ mmol.m}^{-3}$ , the results did not show clear improvements due to the assimilation. There are several reasons that could explain that the assimilation process does not perform well for this range of data. First, these data are generally measured in the oligotrophic

sub-tropical gyre, where the low temporal variability of biogeochemical concentration for the free run (used to compute the EOFs) does not allow the assimilation system to correct strong differences between model and data. Secondly, these data are also generally measured in strong nutrient gradients, i.e. in a transition zone between nutrients enriched waters (high latitudes) and oligotrophic waters of the subtropical gyre. In this region a well-reproduced physics is essential in order to obtain a satisfying biogeochemical dynamics. As the physics was not constrained by the assimilation process, this transition zone may not be located exactly preventing the assimilation system to perform correctly.

## Conclusion

Realistic ocean colour assimilation experiments have been performed in a CPBM of the North Atlantic basin as part of the GMMC Mercator-Vert and FP7 MyOcean projects. These experiments aimed at assessing the feasibility of multi-year reanalyses of the ocean biogeochemistry based on SeaWiFS ocean colour data in a  $\frac{1}{4}^\circ$  eddy-admitting model. Different simulations were conducted during a 9 year-long period (1998-2007), allowing the assimilation of 410 SeaWiFS-estimated maps of chlorophyll temporally binned every 8 days. The assimilation was performed with a fixed variant of the reduced-rank Kalman filter (SEEK) in order to limit the computational burden of the data integration process. Several key parameters entering the analysis step (e.g. model and observation errors parameterizations, local influence radius) were carefully tuned to maximize the benefit of the assimilation process. Comparisons were made between a free run, an assimilation run using a linear updating scheme, and an assimilation run using a non-linear updating scheme with anamorphic transformations.

Overall, the assimilation experiments demonstrate that the SeaWiFS data provide a very powerful means of controlling the seasonal evolution of primary production in the North Atlantic ocean. The basin-averaged surface chlorophyll concentration was computed over several regional domains inside the modelled area. The results showed that the use of a non-linear analysis scheme largely improved the description of this variable along the simulation period. The link between data spatial coverage and assimilation efficiency was underlined for high latitude regions. Nevertheless, unrealistic corrections were detected locally in case where a linear scheme was used to assimilate the data. These anomalies appeared to be the consequence of spurious multivariate corrections of non-observed variables.

The assimilation impact on nitrate variables was assessed using the WOA09 data set. Spatial and temporal colocations were performed between the model outputs and a total of 4110 *in situ* data during the simulation period. The RMS misfit between the assimilation products and unassimilated *in situ* measurements was computed within several vertical intervals of the water column. It appears that the description of the nitrate concentration was improved between the surface and 50 m depth when using the non-linear assimilation scheme. Compared to the free run, the linear scheme improved the surface nitrate values too but a trend to overestimation was observed between models and data in that case. Differences between the free run and the non-linear scheme in the 50-150 m depth interval were rather small, indicating that the assimilation of surface chlorophyll data might not be sufficient to control the CBPM dynamics over the whole euphotic layer. Concerning the deeper part of the water column, the use of a linear scheme clearly increased the nitrate RMS, while the non-linear scheme did not show such a negative impact.

This study opens new perspectives for large-scale biogeochemical simulations assimilating ocean colour data. Firstly, the pre-operational system used here demonstrates that the method is efficient and stable, paving the way for robust real-time biogeochemical nowcast/forecast in the near future. Secondly, the computation of more realistic biogeochemical climatologies and open-sea boundary conditions can be expected from this kind of multi-annual reanalyses. This is a much expected opportunity since the currently available nutrient climatologies based on objective data interpolation are still hampered in a number of ways by the lack of *in situ* data (Garcia *et al.*, 2006).

A number of improvements of the system presented in this article will have to be explored before undertaking operational transition into the Mercator-Ocean suite. As already stated before, the physical simulation was not constrained by any data assimilation scheme in the experiment described in this paper. Constraining the physical variables should improve both the biogeochemical dynamics but also the ocean colour data assimilation efficiency through a better description of the error sources. Currently, an operational simulation is run by the Mercator-Ocean group where the physics is constrained by an assimilation process while the biogeochemical variables evolve freely (so-called BIOMER1V1). Merging the BIOMER1V1 system with the ocean colour data assimilation system described here represents the future of the global ocean biogeochemistry forecasting.

Nevertheless some extensive studies should be performed prior to reach this milestone with regard to the biogeochemical assimilation process. Sensitivity experiments should be intensified to refine the parameterization of model errors (e.g. through ensemble-based approaches) and observation errors, as well as for improving the localization setup. An obvious limitation considered in the present approach arises from the assumption that satellite-derived chlorophyll concentrations are confined to the surface level of the CBPM, while the satellite measured values are actually integrating a finite portion of the euphotic depth. Therefore, considering the verticality of ocean colour measurements could be a more sensible way for combining ocean colour data with biogeochemical models.

## Acknowledgement

This work was conducted as a contribution to the MyOcean Project funded by the E.U. (Grant FP7-SPACE-2007-1-CT-218812-MYOCEAN), with

additional support from the Groupe Mission Mercator Coriolis (GMMC) as part of the Mercator-Vert project. The calculations were performed using HPC resources from GENCI-IDRIS (Grant 011279).

## References

- Barnier, B., Madec, G., Penduff, T., Molines, J.M., Treguier, A.M., Le Sommer, J., Beckmann, A., Biastoch, A., Boning, C., Dengg, J., Derval, C., Durand, E., Gulev, S., Remy, E., Talandier, C., Theetten, S., Maltrud, M., McClean, J., and De Cuevas, B., 2006. Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution. *Ocean Dynamics*, 56: 543–567. doi: 10.1007/s10236-006-0082-1.
- Boyer, T., Antonov, J., Garcia, H., Johnson, D., Locarnini, R., Mishonov, A., Pitcher, M., Baranova, O., and Smolyar, I., 2006. World Ocean Database 2005. U.S. Government Printing Office, Washington D.C., 190 pp.
- Brankart J.-M., Testut C.-E., Béal D., Doron M., Fontana C. Meinvielle M. and Brasseur P., 2011 : Towards an improved description of oceanographic uncertainties : effect of local anamorphic transformations on spatial correlations, *submitted to Ocean Sci.*
- Brasseur P. and Verron J., 2006 : The SEEK filter method for data assimilation in oceanography: a synthesis, *Ocean Dynamics*, 56, 650-661, doi :10.1007/s10236-006-0080-3.
- Brasseur, P., Gruber, N., Barciela, R., Brander, K., Doron, M., El Moussaoui, A., Hobday, A.J., Huret, M., Kremer, A.S., Lehodey, P., Matear, R., Moulin, C., Murtugudde, R., Senina, I., and Svendsen, E., 2009. Integrating Biogeochemistry and Ecology Into Ocean Data Assimilation Systems. *Oceanography*, 22(3): 206–215.
- Béal, D., Brasseur, P., Brankart, J.M., Ourmières, Y., and Verron, J., 2010. Characterization of mixing errors in a coupled physical biogeochemical model of the North Atlantic : implications for nonlinear estimation using Gaussian anamorphosis. *Ocean Science*, 6(1): 247–262. doi :10.5194/os-6-247-2010.
- Carmillet, V., Brankart, J.M., Brasseur, P., Drange, H., Evensen, G., and Verron, J., 2001. A singular evolutive extended Kalman filter to assimilate ocean colour data in a coupled physical-biochemical model of the North Atlantic ocean. *Ocean Modelling*, 3(3-4) : 167 – 192. doi: 10.1016/S1463-5003(01)00007-5.
- Doron, M., Brasseur, P., and Brankart, J.M., 2011. Stochastic estimation of biogeochemical parameters of a 3D ocean coupled physical-biochemical model : Twin experiments. *J. Mar. Sys.*, 87(3-4): 194 – 207. doi: 10.1016/j.jmarsys.2011.04.001.
- Feldman, G., McCain, C., 2010. Ocean colour web, seawifs reprocessing 2009.1. NASA Goddard Space Flight Center. Eds. Kuring, N., Bailey, S. W., <http://oceancolor.gsfc.nasa.gov/>.
- Fontana, C., Grenz, C., Pinazo, C., Marsaleix, P., and Diaz, F., 2009. Assimilation of SeaWiFS chlorophyll data into a 3D-coupled physical biogeochemical model applied to a freshwater-influenced coastal zone. *Continental Shelf Research*, 29(11-12) : 1397 – 1409. doi: 10.1016/j.csr.2009.03.005.
- Fontana, C., Grenz, C., and Pinazo, C., 2010. Sequential assimilation of a year-long time-series of SeaWiFS chlorophyll data into a 3D biogeochemical model on the French Mediterranean coast. *Continental Shelf Research*, 30(16): 1761 – 1771. doi: 10.1016/j.csr.2010.08.003.
- Garcia, H., and T.P. Boyer, R.L., and Antonov, J., 2006. World Ocean Atlas 2005, Volume 4 : Nutrients (phosphate, nitrate, and silicate). U.S. Government Printing Office, Washington D.C., 396 pp.
- Gregg, W.W. and Casey, N.W., 2004. Global and regional evaluation of the SeaWiFS chlorophyll data set. *Remote Sensing of Environment*, 93 (4): 463 – 479. doi: 10.1016/j.rse.2003.12.012.
- Lavender, S., Pinkerton, M., Moore, G., Aiken, J., and Blondeau-Patissier, D., 2005. Modification to the atmospheric correction of SeaWiFS ocean colour images over turbid waters. *Continental Shelf Research*, 25(4): 539 – 555. doi: 10.1016/j.csr.2004.10.007.
- Longhurst, A., 1995. Seasonal cycles of pelagic production and consumption. *Progress In Oceanography*, 36(2): 77 – 167. doi: 10.1016/0079-6611(95)00015-1.
- Lévy, M., Gavart, M., Mémerly, L., Caniaux, G., and Paci, A., 2005. A four-dimensional mesoscale map of the spring bloom in the northeast Atlantic (POMME experiment): Results of a prognostic model. *J. Geophys. Res.*, 110(C7): C07S21.

Natvik, L.J. and Evensen, G., 2003. Assimilation of ocean colour data into a biochemical model of the North Atlantic : Part 1. Data assimilation experiments. *Journal of Marine Systems*, 40-41: 127 – 153. doi: 10.1016/S0924-7963(03)00016-2. The Use of Data Assimilation in Coupled Hydrodynamic, Ecological and Bio-geo-chemical Models of the Ocean. Selected papers from the 33rd International Liege Colloquium on Ocean Dynamics, held in Liege, Belgium on May 7-11th, 2001.

Nerger, L. and Gregg, W.W., 2007. Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. *Journal of Marine Systems*, 68(1-2): 237 – 254. doi: 10.1016/j.jmarsys.2006.11.009.

Nerger, L. and Gregg, W.W., 2008. Improving assimilation of SeaWiFS data by the application of bias correction with a local SEIK filter. *Journal of Marine Systems*, 73(1-2): 87 – 102. doi: 10.1016/j.jmarsys.2007.09.007.

Ourmières, Y., Brasseur, P., Lévy, M., Brankart, J.M., and Verron, J., 2009. On the key role of nutrient data to constrain a coupled physical-biogeochemical assimilative model of the North Atlantic Ocean. *Journal of Marine Systems*, 75(1-2): 100 – 115. doi:10.1016/j.jmarsys.2008.08.003.

Pham, D.T., Verron, J., and Roubaud, M.C., 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, 16(3-4): 323 – 340. doi: 10.1016/S0924-7963(97)00109-7.

Schiller, A., Wijffels, S.E., and Meyers, G.A., 2004. Design Requirements for an Argo Float Array in the Indian Ocean Inferred from Observing System Simulation Experiments. *Journal of Atmospheric and Oceanic Technology*, 21(10): 1598–1620. doi: 10.1175/1520-0426(2004)021.

# IFREMER/CERSAT ARCTIC SEA ICE DRIFT 1992-2010 TIME SERIES FROM SATELLITE MEASUREMENTS FOR MYOCEAN

By **Fanny Girard-Arhuin<sup>(1)</sup>**, **Denis Croize-Fillon<sup>(2)</sup>**

<sup>(1)</sup>IFREMER, Laboratoire d'Océanographie Spatiale, Plouzané, France

<sup>(2)</sup>IFREMER, CERSAT, Plouzané, France

## Abstract

Satellites enable daily and global coverage of the polar oceans, providing a unique monitoring capability of sea ice dynamics. Sea ice drift fields can be estimated in Arctic from several sensors, in particular scatterometers and radiometers. The 1992-2010 winter sea ice drift time series made at Ifremer/CERSAT and available for the GMES-MyOcean project (<http://www.myocean.eu>) are presented here. There are several products : the Level 3 products from single sensors and the L4 products from the combination of sensors. They are available at 3, 6 and 30 day-lag with a 62.5 km-grid size during winter. Results of the validation with in situ data are shown. Drift data density shows that almost full maps are available over the winter thanks to the merging of the data and the interpolation of the missing points. These data are available for operational use and the scientific community.

## Introduction

Sea ice cover and motion have major impacts on heat fluxes between ocean and atmosphere in polar areas. Moreover, the Arctic ice growth and melt impact the fresh water flux which then has an important role in the thermohaline circulation, in particular in European seas (Aagaard et al., 1985). This is why sea ice observations are required, in particular long time series in order to better understand the role of sea ice in ocean circulation or climate change.

The sea ice drifts shown here concern only the Arctic ocean where it has been extensively validated with in situ data. Since 1979, buoys are moored in Arctic ice, each year, in the framework of the International Arctic Buoy Program (IABP). They provide continuous local measurements (pressure, air temperature, ice drift) covering mostly the Central Arctic and Beaufort Sea, but the spatial distribution is rather sparse in the Eurasian Seas, providing a limited view of the velocity field (Colony and Thorndike, 1984). Since the 1990's, sea ice drift can be estimated from satellite data, daily and with global coverage of the polar ocean. Given the sensor spatial resolution and the magnitude of the expected drifts, the signature of the feature to be tracked must persist for several days. The ice motion is thus estimated at the scale of large ice floes, whereas a buoy measures the drift of a single ice floe. Data from high spatial resolution satellite sensors have been used for sea ice drift estimate in feasibility studies, for example with Advanced Very High Resolution Radiometer (AVHRR, Emery et al., 1991) and Synthetic Aperture Radar (Kwok et al., 1990). These demonstrations are limited to local or regional studies. Here, we focus on sea ice drift at large scale (or polar scale). These estimates are inferred from passive (radiometers) and active microwave sensors (scatterometers). High frequencies data are very sensitive to atmospheric effects (moisture and liquid water). This implies sea ice drift estimations at periods of low water vapour during winter, from October to April (for example Martin and Augstein, 2000).

This paper presents some Arctic sea ice drift datasets inferred from satellite measurements; they are produced daily in winter at Ifremer via the Centre d'Exploitation et de Recherche SATellitaire (CERSAT), and are available as a time series for the MyOcean project. The datasets presented here are the level 3 products corresponding to the daily Arctic drift data at 3 and 6 day-lags from single sensors, and the level 4 products corresponding to the merging of radiometer and scatterometers sea ice drift fields. These time series cover the period 1992-2010 and are available from October until April (cold period). Section 1 presents the satellite sensors used, Section 2 the sea ice drift fields (level 3 products), Section 3 the merging of sea ice drift fields (level 4 products). Applications of these datasets are in Section 4, conclusions follow in Section 5.

## Sensors

In this section, we present briefly the passive and active microwave sensors onboard satellites we use to infer the sea ice drift fields.

## Radiometers

Since the 1980's, passive microwave radiometers onboard satellites as the Special Sensor Microwave Imager(s) (SSM/I) have been commonly used to estimate sea ice concentration from the daily brightness temperature data. The pixel resolution is 25 km at the lower frequencies and 12.5 km with the 85.5 GHz channel, we focus here on the 12.5 km resolution data which starts in 1992.

Passive microwave radiometers data like SSM/I brightness temperature data are commonly used to estimate sea ice concentration, they also have been widely used to estimate Arctic sea ice drift. In order to determine sea ice drift from successive brightness temperature maps, stable radiation of the sea ice cover and negligible influence of the atmospheric conditions are needed.

## Scatterometers

The benefit of the all-weather, day-night microwave radar measurements has been well established. During the 1999-2009 period, daily averaged backscatter maps at the resolution of 12.5 km have been built from the Ku-band SeaWinds/QuikSCAT scatterometer data. Since 2007, ASCAT scatterometer onboard MetOp provides C-band backscatter data we process to have daily averaged backscatter maps (same pixel size resolution). These data are used to infer sea ice drift fields.

## Single sensors sea ice drift fields

Sea ice drift fields at global scale exist since the beginning of the Arctic monitoring with the low frequency channels of the SSM/I radiometer sensor. In this section, some methods to infer sea ice drifts are briefly presented, and the one applied at Ifremer/CERSAT highlighted. Results of validation with in situ buoys data follow.

### Methods

All methods assume that the structures tracked have spatial dimensions larger than the pixel resolution. For the pixel resolution available, single ice floes cannot be detected. Several methods have been tested to determine sea ice displacements : algorithms are based on tracking common features in pairs of sequential satellite maps.

One tool which is the most widely used is the Maximum Cross Correlation (MCC), it was used successfully with SSM/I or AVHRR data for example (Emery et al., 1991; Ninnis et al., 1986). This method enables detection of translation displacement and no rotation (Kamachi, 1989; Ninnis et al., 1986). A correlation is estimated at each position between two arrays: one at a given day and another one lagged in time in order to obtain a correlation coefficient array. The location of the maximum correlation is the location of the maximum similarity between the two original images. The displacement can thus be inferred. Details about tracking methods can be found among others in Ezraty et al. (2007a,b), Martin and Augstein (2000), Emery et al. (1997), Kwok et al. (1998), and in Lavergne et al. (2010) who introduce recently an alternative tracking method based on MCC.

For the products presented here, we apply the MCC tool to the Laplacian field of brightness temperature or backscatter data (from radiometer and scatterometer sensor respectively) in order to enhance the structures to be tracked. In order to remove outliers, a minimum coefficient correlation is imposed. In addition, a comparison with the wind pattern is applied (here ECMWF model wind field, NCEP re-analyses from Kwok et al., 1998) since mean sea ice drift is strongly linked with the geostrophic wind (Thorndike and Colony, 1982).

Another tool to enhance the feature to be tracked is the wavelet analysis. It is applied on various spatial scales in order to separate ice features. A two-dimensional Gaussian wavelet (also called Mexican hat wavelet) has been used for ice feature detection by Liu and Cavalieri (1998) and Liu et al. (1999). It turns out that this wavelet technique is very similar to the Laplacian field enhancement as used for the Ifremer/CERSAT product.

### Results

Some results of SSM/I drift validation are summarized in Table 1. The displacements (in km) are usually converted into ice speed (in  $\text{cm s}^{-1}$ ) in order to compare the quality of the drift magnitude estimated at different day-lags or at various pixel size. Nevertheless, this introduces a non linear effect in the ratio. Moreover, the constant half-pixel minimum displacement embedded in the original measurement is then erroneously accounted for.

These results show that with SSM/I, sea ice drifts can be retrieved with good accuracy. Note that with a pixel resolution of 12.5 km, for 1 day-lag, the minimum displacement that can be detected is 6.25 km which corresponds to a velocity of  $7 \text{ cm s}^{-1}$ , whereas the mean buoys velocity are about  $7 \text{ cm s}^{-1}$ . Then the day-lag value to be used for a 12.5 km resolution pixel must be larger: 3 day-lag corresponds to a minimum velocity of  $4.8 \text{ cm s}^{-1}$ . The 3-day lag is commonly used. The day-lag must be chosen accordingly to the magnitude of the expected drift (small drift – large

lag; large drift – short lag). That's why the Ifremer/CERSAT products used 3 day-lag and 6 day-lag to infer large and small drifts (Ezraty et al., 2007a). The time series of these drifts inferred from SSM/I channels are available since 1992. A 2 day-lag is also used with lower resolution pixel data like the AMSR-E radiometer data (Ezraty et al., 2007c, Laverigne et al., 2010).

std speed difference ( $\text{cm s}^{-1}$ )	std direction difference ( $^{\circ}$ )	day-lag (nb of days)	reference
6		1	Emery et al., 1997
2.6 à 2.9	18 à 25.9	1	Liu and Cavalieri, 1998
4.28 → 4.52 2.58 * → 2.89 *	46.5 → 50.4 30.6 * → 32.8 *	3	Kwok et al., 1998
2.96	34.4	4	Liu et al., 1999
2.27 *	35.5 *	4	Zhao et al., 2002

*Table 1—Results of the comparison between SSM/I drift vectors and buoys drifts for several experiments : standard deviation (std) for sea ice drift difference and for direction difference, number of day-lag, and reference. \* indicates that the drift magnitudes less than one pixel are excluded.*

The SeaWinds/QuikSCAT and the ASCAT/MetOp scatterometer daily averaged backscatter maps are also used at Ifremer/CERSAT to estimate sea ice drift fields applying the same method at 3 and 6 day-lag (Ezraty et al., 2007b). The SeaWinds/QuikSCAT drift field time series is available for the period 1999-2009, the ASCAT/MetOp time series has started in 2007 and is ongoing. Figure 1 shows the drift data density during the 2003-2004 winter. Radiometer channels (SSM/I H & V) provides 60 to 75% data density during the winter, less at the beginning of the winter, which is the worst period of estimation. QuikSCAT data density has a slightly higher density, in particular during the January-April period with a density reaching about 80%.

Sea ice drifts deduced from radiometers have a reasonable accuracy but are limited by data gaps and low data density at the beginning and sometimes also at the end of the cold period. This can be improved using scatterometer data to build a merged product.

std speed difference ( $\text{cm s}^{-1}$ )	std direction difference ( $^{\circ}$ )	day-lag (nb of days)
2.91	39.2	3
2.90 *	29.6 *	
1.72	29.6	6
1.76 *	24.3 *	

*Table 2— Results of the comparison between Ifremer/CERSAT "Merged" SSM/I/QuikSCAT drift vectors and buoys drifts for five winters : standard deviation for sea ice drift difference and for direction difference, and number of day-lag. \* indicates that the drift magnitudes less than one pixel are excluded.*

## Merging sea ice drift fields

Due to satellite geometry and near-polar orbit, the data gap near the North Pole for the scatterometers is smaller than that of the SSM/I. Moreover, radars at Ku-band do not have the problem caused by water vapour as the 85.5 GHz SSM/I channels, so QuikSCAT is an adequate sensor to fill in poorly tracked areas where the weather influence prevents reliable SSM/I estimation during the cold season. We present here the advantages brought by the combination of several independent drift fields from radiometer and scatterometer as performed at Ifremer/CERSAT, more details can be found in Girard-Ardhuin et al. (2008).

Sea ice drift maps from SSM/I are based on the two channels of the sensor (horizontal and vertical polarizations, 'H' and 'V'): from each channel, a drift map is inferred. This results in two independent drift fields. Merging the two fields gives a drift product which is more reliable than the one inferred from a single channel. We also combine these two fields and the QuikSCAT sea ice drift field by optimal merging, leading to the "Merged" product.

The Merged product has been validated using IABP buoys during five winters (see Table 2): the standard deviation of the difference at 3 day-lag is  $2.91 \text{ cm s}^{-1}$ , comparable with previous results quoted above for 3 and 4 day-lags (Table 1). Using 6 day-lag is even more adequate for small

drifts with a standard deviation of the difference of  $1.72 \text{ cm s}^{-1}$ . Consequently, the angular data have a better resolution. The angle difference sharply decreases (to values lower than  $45^\circ$ ) for drifts higher than 40 km (also noticed by Liu et al., 1999). Standard deviations decrease when small drifts (displacement lower than the pixel size) are excluded, in particular for drift direction.

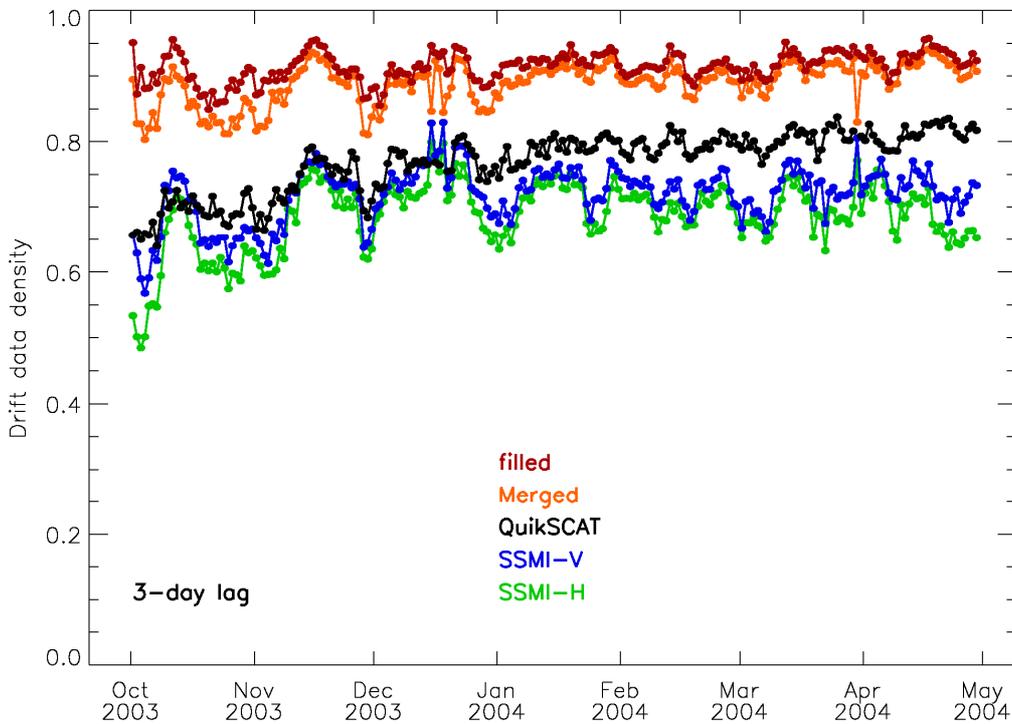


Figure 1 : Times series of drift data density (percentage x 100) for the 2003-2004 winter for SSM/I at horizontal (in green) and vertical (in blue) polarizations, QuikSCAT (in black), Merged SSM/I/QuikSCAT (in red) and filled Merged SSM/I/QuikSCAT (in brown) drifts at 3 day-lag.

The combination of the independent drift fields provides better confidence in the final resulting field than for the individual ones since each drift is inferred from independent measurement. First, the data gap at the North Pole is reduced to the smallest one. Second, the sea ice Merged drift field enables discrimination of remaining vectors outliers of the individual products. Third, the number of valid drift vectors is higher than the single sensors valid drift vectors: the merging increases the data density up to 90% for December-April period and more than 80% at fall in the example of figure 1.

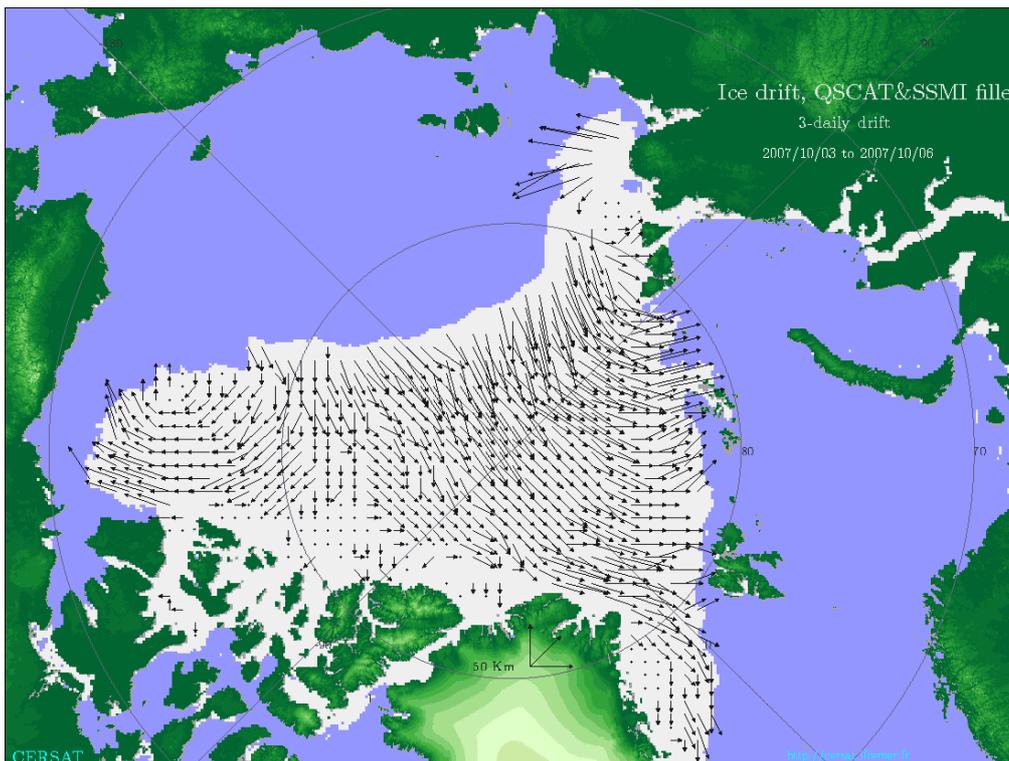


Figure 2: Merged filled Arctic sea ice drift field from SSM/I radiometer drift fields and Sea-Winds/QuikSCAT scatterometer drift field at 3 day-lag on October 3<sup>rd</sup>-6<sup>th</sup>, 2007. Grid spacing is 62.5 km. Drift vectors less than one pixel are marked with a cross.

In order to ease the use of these datasets by modelers, we have developed a space and time interpolation to infer drift in some data gaps patches. Figure 2 shows an example of this filled drift field at 3 day-lag in October, at the beginning of freeze period, with an almost full drift map, thanks to the interpolation which enables to increase by 5 to 10% the drift data density (also see in Fig. 1 in brown). This space and time interpolation is very useful, in particular for fall estimation when drift is more difficult to infer.

## Applications

Drift datasets are useful for several applications. Among others, Ifremer/CERSAT datasets have been used to estimate pollution origin location (H. Goodwin from the Norwegian Polar Institute, *personal comm.*), to study sea ice drift (Spren et al., 2011), and to estimate of sea ice flux in Fram Strait or Baffin Bay (for example Spren et al., 2006; Spren et al., 2009; Kwok, 2007).

They are also commonly used in models, in particular for validation (Martin and Gerdes, 2007 and recently Rozmann et al., 2011). They are also used in operational models for assimilation. Tests have been done in the UK MetOffice FOAM model showing better realistic results for sea ice motion when the Ifremer/CERSAT products are assimilated (Stark et al., 2008). Assimilation of these fields greatly improves results on the modelled drift but also on sea ice concentration and thickness for Rollenhagen et al. (2009). Similarly, the assimilation in the TOPAZ systems enables the modelled sea ice thickness to be more reliable (Bertino et al., 2011).

## Conclusions

Satellites provide a unique monitoring capability of sea ice dynamics. Sea ice maps from satellite data have the advantage to cover the whole polar areas every day. This allows better coverage than that of buoys data, which have a precise daily timing, better spatial resolution but a very coarse spatial distribution.

Sea ice drifts inferred only from radiometers have a reasonable accuracy but are limited by data gaps and low data density at the beginning and the end of the cold period. The optimal merging of independent fields of drift data (SSM/I fields and scatterometer fields from QuikSCAT or ASCAT) at the same resolution improves the data density and the usable time period over winter. It also enables the discrimination of the vector outliers remaining in the individual products.

The scatterometers (QuikSCAT/ASCAT) drift fields and filled Merged (SSM/I/QuikSCAT and ASCAT/SSM/I) Arctic sea ice drift fields datasets are available every day during winter (October until April). The filled merged fields are available since 1992, using SSM/I horizontal and vertical polarizations data from 1992 until 1999 and merging with QuikSCAT data for the 1999-2009 period, ASCAT is merged with SSM/I data since 2007. Monthly drifts are also computed. The interpolated merged field has the advantage to be gap-filled motion fields ready-to-use.

Ifremer/CERSAT hosts a unique database of 18 years winter time series of sea ice drift. This dataset is available for oceanic and climate modelling and various scientific studies in the Arctic. The time series is ongoing and will continue for Arctic long term monitoring using the next MetOp/ASCAT operational scatterometers, planned to be operated for the next 10 years.

## Acknowledgement

The authors thank Robert Ezraty and Julien Paul for their work on the merging and the space and time interpolation developments. They also thank the CERSAT team for acquisition and archiving of the data and the systematic processing of the sea ice drift maps. SSM/I data are provided by National Snow and Ice Data Center, University of Colorado (USA), QuikSCAT data by Jet Propulsion Laboratory (USA), and ASCAT raw data by EUMETSAT. The International Arctic Buoy Programme buoys data are provided by the Polar Science Center, University of Washington (USA). The Global Monitoring for Environment and Security European Union projects PolarView and MyOcean support partly the data production and distribution.

## References

- Aagaard, K., Swift, J. H. and Carmack, E. C. 1985: Thermohaline circulation in the Arctic Mediterranean seas. *J. Geophys. Res.* 90 (C3): 4833–4846.
- Bertino, L., Sakov, P., and Counillon F. 2011: The TOPAZ sea ice and ocean pilot reanalysis (2003-2008). European Geophysical Union General Assembly meeting, 3-8 April 2011, Vienna, Austria.

- Colony, R. and Thorndike, S. 1984: An estimate of the mean field of Arctic sea ice motion. *J. Geophys. Res.* 89 (C6): 10623–10629.
- Emery, W. J., Fowler, C. W., Hawkins, J. and Preller, R. H. 1991: Fram strait satellite image-derived ice motions. *J. Geophys. Res.* 96 (C3): 4751–4768.
- Emery, W. J., Fowler, C. W. and Maslanik, J. A. 1997: Satellite-derived maps of Arctic and Antarctic sea ice motion: 1988 to 1994. *Geophys. Res. Lett.* 24 (8): 897–900.
- Ezraty, R., Girard-Ardhuin, F., Piollé, J. F., Kaleschke, L. and Heygster, G. 2007a: Arctic and Antarctic sea ice concentration and Arctic sea ice drift estimated from Special Sensor Microwave data, User's manual, version 2.1, February 2007. Available at IFREMER/CERSAT.
- Ezraty, R., Girard-Ardhuin, F. and Piollé, J. F. 2007b: Sea ice drift in the central Arctic combining QuikSCAT and SSM/I sea ice drift data, User's manual, version 2.0, February 2007. Available at IFREMER/CERSAT.
- Ezraty, R., Girard-Ardhuin, F. and Croizé-Fillon, D. 2007c: Sea-ice drift in the central Arctic using the 89 GHz brightness temperatures of the Advanced Microwave Scanning Radiometer, User's manual, version 2.0, February 2007. Available at IFREMER/CERSAT.
- Girard-Ardhuin, F., Ezraty, R., Croizé-Fillon, D., and Piollé, J.F. 2008: Sea-ice drift in the central Arctic combining QuikSCAT and SSM/I sea ice drift data, User's manual, version 3.0, April 2008. Available at IFREMER/CERSAT.
- Kamachi, M., 1989: Advective surface velocities derived from sequential images for rotational flow field : limitations and applications of Maximum Cross Correlation method with rotational registration. *J. Geophys. Res.* 94 (C12): 18227–18233.
- Kwok, R., Curlander, J.C., McConnell, R. and Pang, S. S. 1990: An ice motion tracking system at the Alaska SAR Facility. *IEEE. J. Ocean. Eng.* 15 (1): 44–54.
- Kwok, R., Schweiger, A., Rothrock, D. A., Pang, S. and Kottmeier, C. 1998: Sea ice motion from satellite passive microwave imagery assessed with ERS SAR and buoy motions. *J. Geophys. Res.* 103 (C4): 8191–8214.
- Kwok, R. 2007: Baffin bay ice drift and export: 2002-2007. *Geophys. Res. Lett.* 34, doi:10.1029/2007GL031204.
- Lavergne, T., Eastwood, S., Teffah, Z., Schyberg, H. and Breivik, L.A. 2010: Sea ice motion from low resolution satellite sensors: an alternative method and its validation in the Arctic. *J. Geophys. Res.* 115, doi:10.1029/2009JC005958.
- Liu, A. K. and Cavalieri, D. J. 1998: On sea ice drift from the wavelet analysis of the Defense Meteorological Satellite Program (DMSP) Special Sensor Microwave Imager (SSM/I) data. *Int. J. Rem. Sens.* 19 (7): 1415–1423.
- Liu, A. K., Zhao Y. and Wu, S. Y., 1999: Arctic sea ice drift from wavelet analysis of NSCAT and special sensor microwave imager data. *J. Geophys. Res.* 104 (C5): 11529–11538.
- Martin, T. and Augstein, E. 2000: Large-scale drift of Arctic sea ice retrieved from passive microwave satellite data. *J. Geophys. Res.* 105 (C4): 8775–8788.
- Martin T., and Gerdes, R. 2007: Sea ice drift variability in Arctic ocean model intercomparison project models and observations. *J. Geophys. Res.* 112. doi:10.1029/2006JC003617.
- Ninnis, R. M., Emery, W. J. and Collins, M. J. 1986: Automated extraction of pack ice motion from Advanced Very High Resolution Radiometer imagery. *J. Geophys. Res.* 91 (C9): 10725–10734.
- Rollenhagen, K., Timmermann, R., Janjic, T., Schröter, J. and Danilov, S. 2009: Assimilation of sea ice motion in a finite-element sea ice model. *J. Geophys. Res.* 114. doi:10.1029/2008JC005067.
- Rozmann, P., Hölemann, J., Krumpen, T., Gerdes, R., Köberle, C., Lavergne, T., Adams, S., and Girard-Ardhuin, F. 2011: Validating satellite derived and modeled sea ice drift in the Laptev sea with In Situ measurements of winter 2007/08. *Polar Research.* 30 (7218). doi: 10.3402/polar.v30i0.7218.
- Spreen, G., Kern, S., Stammer, D., Forsberg, R. and Haarpaintner, J. 2006: Satellite-based estimates of sea ice volume flux through Fram strait. *Annals of Glaciology.* 44: 321-328.
-

Spren, G., Kern, S., Stammer, D., and Hansen, E. 2009: Fram strait sea ice volume export estimated between 2003 and 2008 from satellite data. *Geophys. Res. Lett.* 36. doi:10.1029/2009GL039591.

Spren, G., Kwok, R., Menemenlis, D. 2011: Trends in Arctic sea ice drift and role of wind forcing: 1992-2009. *Geophys. Res. Lett.* doi:10.1029/2011GL048970, in press.

Stark, J., Ridley, J., Martin, M. and Hines, A. 2008: Sea ice concentration and motion assimilation in a sea ice-ocean model. *J. Geophys. Res.* 113. doi :10.1029/2007JC004224.

Thorndike, A. S. and Colony, R. 1982: Sea ice motion in response to geostrophic winds. *J. Geophys. Res.* 87: 5845–5852.

Zhao, Y., Liu A. K. and Long, D. G. 2002: Validation of sea ice motion from QuikSCAT with those from SSM/I and buoy. *IEEE Trans. On Geosc. and Rem. Sens.* 40 (6): 1241–1246.

## NOTEBOOK

### Editorial Board

Laurence Crosnier

### DTP Operator

Fabrice Messal

### Articles

**News: Workshop on Observing System Evaluation and Intercomparisons : a GODAE Ocean view/GSOP/CLIVAR event (13-17 June 2011, Santa Cruz, USA)**

**CORA3, a comprehensive and qualified ocean in-situ dataset from 1990 to 2010**

*By Cécile Cabanes, Antoine Grouazel, Victor Turpin, François Paris, Christine Coatanoan, Karina Von Schuckmann, Loic Petit de la Villeon, Thierry Carval, Sylvie Pouliquen*

**A multivariate reanalysis of the North Atlantic ocean biogeochemistry during 1998-2007 based on the assimilation of SeaWiFS data**

*By Clément Fontana, Pierre Brasseur, Jean-Michel Brankart*

**Ifremer/CERSAT Arctic sea ice drift 1992-2010 time series from satellite measurements for myocean**

*By Fanny Girard-Arhuin, Denis Croize-Fillon*

### Contact :

Please send us your comments to the following e-mail address: [webmaster@mercator-ocean.fr](mailto:webmaster@mercator-ocean.fr).

**Next issue : January 2012**

---