

# SANGOMA: Stochastic Assimilation for the Next Generation Ocean Model Applications EU FP7 SPACE-2011-1 project 283580

## Final report



J.-M. Beckers    A. Barth    Y. Yan    L. Vandenbulcke  
M. Canter    F. Laenen  
University of Liège, BELGIUM

P.-J. Van Leeuwen    S. Vetra-Carvalho  
University of Reading, UK

L. Nerger    P. Kirchgessner  
Alfred-Wegener-Institut, GERMANY

A. Heemink    N. van Velzen    M. Verlaan    U. Altaf  
Delft University of Technology, NETHERLANDS

P. Brasseur    J.-M. Brankart    G. Candille    S. Metref  
F. Garnier and A. Bichet  
CNRS-LEGI, FRANCE

P. De Mey    N. Ayoub  
CNRS-LEGOS, FRANCE

L. Bertino    F. Counillon    A. Carrassi  
NERSC, NORWAY

# Chapter 1

## Summary

### 1.1 How to use this document

The present document is meant to provide a quick overview of the main achievements of the SANGOMA project and is not intended to include all relevant or detailed information from the individual reports on deliverables. It contains pointers to the most important deliverables for further reading.

### 1.2 Project objectives

The objectives and expected impacts from the project description did not change during the project evolution and read

SANGOMA will provide new developments in data assimilation to ensure that future operational systems make use of state-of-the-art data-assimilation and related analysis tools. We are a European network of expert teams in advanced data assimilation. In the project we will extend existing modular data assimilation systems that have high flexibility in type of ocean model and assimilation method. Following specific design rules, new modules can be used in different modular systems. The systems will allow for efficient operational testing of the latest data assimilation methods, and quick comparison of assimilation methods for operational use. Furthermore, we will develop and implement modules that objectively determine the impact of existing and new observation types. This dedicated web portal will provide access to validated products, including documented performances on a variety of test cases. Consolidated versions will be made available to the science community and Marine Forecasting Centres with indications on best practice implementation. Workshops and summer schools on advanced assimilation methods and modular systems will ensure fast and efficient training of next generation oceanographers, ensuring world-leading operational oceanographic products for costumers and decision makers.

The developments of SANGOMA will also serve costumers of My-Ocean products, which is the first European project dedicated to the

implementation of the GMES Marine Core Service for ocean monitoring and forecasting. For this purpose, we will concentrate on data-assimilation methods that deliver probabilistic information on the products. To this end, existing ensemble methods will be included and new methods that allow for nonlinear and non-Gaussian systems will be developed.

### 1.3 Achievements

- WP1: The planned harmonization of data models and interfaces was not easy *a priori* and an "Keep it simple and short" approach was chosen. The objective has been achieved and, interestingly, the defined standards proven adequate after the fact, including for the advanced EWPF (Equivalent Weight Particle Filter) implementation.
- WP2: A nice collection of tools was elaborated with a complete set of examples. Particularly diagnostic tools with simple and advanced techniques are of wide interest. Also a common framework for a series of variants of EnKF (Ensemble Kalman Filter) was provided as a starting point for newcomers in data assimilation.
- WP3: Nice reviews and presentations of the EnKF variants in a coherent framework and terminology were synthesized. New theoretical methods were developed and tested, from which two (EWPF+Anamorphosis) were implemented into toolboxes with SANGOMA standards.
- WP4: It includes benchmark definitions (taken up by other groups outside of SANGOMA) with new metrics, in particular probabilistic metrics for ensemble approaches.
- WP5: Realistic applications with new data types were implemented and analyzed.
- WP6: Disseminated the results and achieved a strong scientific output.

## Chapter 2

# Objectives and achievements

### 2.1 WP1: Harmonization

This workpackage aimed at standardizing assimilation tools and file formats for easier exchange and use both by project partners and other data assimilation experts. The first year has been focusing on

- analysing the different toolboxes currently in use (content, data model, calling interfaces, etc.),
- defining an initial list of exiting tools from those toolboxes to be shared,
- defining a data model and calling interface for SANGOMA tools.

A crucial point was the definition of data models and calling interfaces. Here, the adopted solution was to keep them simple in order to allow for rapid use of the tools by non specialists in advanced programming methods. Furthermore, the broad range of interfaces and programming languages used in the existing toolboxes made necessary a common denominator, leading to the following:

- for data exchange via files: use of netCDF files in CF compliant form. As the format allows for some variants, the general strategy is to "keep it simple" and to provide output files in a similar form than input files (even if not perfectly fitting CF conditions as is the case for some model outputs). Also version 3 features of netCDF are to be preferred over version 4 to enhance backward compatibility. Finally, ensembles will be treated by working on a collection of files instead of a single big file.
- for data exchange in memory (subroutine call): Use of basic FORTRAN structure *arrays*. No derived types are allowed (too much programming overhead in filling or adapting data types). For more complex interfacing or data structures the *call-back* approach has to be used. Example: if the observational error variance matrix  $\mathbf{R}$  is non-diagonal and a tool needs to evaluate  $\mathbf{R}\mathbf{y}$ , the interface of the tool must include a call-back function which when called with argument  $\mathbf{y}$  returns the product  $\mathbf{R}\mathbf{y}$ . In this case, in the call-back program more complex (system specific) structures can be used without the need to define complicated interfacing in the SANGOMA tools. C-binding specifications are also provided.

```

module sangoma_callback

  use, intrinsic :: ISO_C_BINDING
  use sangoma_base, only:REALPREC, INTPREC
  implicit none

contains

  subroutine some_operation(x, n, f_callback) &
    bind(C,name="callback_some_operation")

    use, intrinsic :: ISO_C_BINDING
    implicit none

    integer(INTPREC), value, intent(in) :: n
    real(REALPREC),          intent(in) :: x(n)

    interface
      subroutine f_callback(x,n) bind(C)
        use, intrinsic :: ISO_C_BINDING
        use sangoma_base, only:REALPREC, INTPREC

        integer(INTPREC), value, intent(in) :: n
        real(REALPREC),          intent(in) :: x(n)
      end subroutine
    end interface

```

The proposed formats and interfaces defined in the first reporting period were well accepted (internally and externally to the project, as proved by the survey) and implementation started. Only minor details regarding the definitions remained to be clarified and were discussed at the second progress meeting:

**F90/F77** : As legacy of F77 is still important it was decided not to force people to use .F90 interfaces and keep things as simple as possible. For Fortran 90 users, it would be useful to provide a SANGOMA module, so that the compiler can automatically check if all mandatory parameters are present and have the correct type. Partners investigated a way to provide a Fortran 77 library and a Fortran 90 module without duplicating the code base.

**Single/double** : some special handling is needed when interfacing with BLAS/LAPACK libraries when different precisions are used. As we cannot serve all situations and as users will generally optimize anyway the codes (memory/speed) when necessary it is decided to focus on the readability of the codes and examples provided with the codes rather than on fully flexible codes with selectable precision. Therefore double precision is retained as reference for SANGOMA.

**Storage ordering** : The memory storage of arrays (as specified in the common data specification) is column-major (which is used in Fortran and Matlab/Octave). Row-major languages such as C and Java need to change the order of the indices. For performance reason, no explicit memory reordering is performed. The adopted approach is equivalent to the one used in the NetCDF library where the order of the dimensions in a C program is reversed compared to the order of dimensions in a Fortran program.

The specifications have been presented in the MyOcean Science days followed by an on-line survey <http://www.surveymonkey.com/s/ZX3P9D8>. The specifications found a large positive feedback in the DA community as well as the advisory board.

The main deliverable of WP1 is therefore DL1.3. The other deliverables reflect the process of tool selections for inclusion in the common toolbox during the

project life. The final outcome of this process is highlighted in the tool descriptions of WP2.

### 2.1.1 Deviations from DOW

None

#### Highlight:

The partners found a common approach despite using quite different data models at their institutes. The adopted internal standard did not need a revision during the project live and also allowed to encompass the case of the advanced Data assimilation tool (EWPF) coupled to the different DA toolboxes using the defined SANGOMA standards and interfaces. For technical details see DL1.3.

## 2.2 WP2: Sharing and collaborative development

This workpackage aimed at sharing tools in a collaborative way using standards set up in WP1.

Technically, the sharing was done via a server with version control. The SANGOMA project was registered on <http://sourceforge.net/projects/sangoma/> as it allows not only sharing of files but also easy setup of discussion forums, download procedures etc. All SANGOMA scientists have registered and are allowed to upload changes (commit of software changes). The site not only contains assimilation-related software but also `html` files for the web server and document templates so that each partner was able to easily contribute.

Within WP2 of SANGOMA, a collection of tools of general interest for data assimilation applications was prepared. These tools provide additional functionality outside of the core of the different tool boxes for data assimilation.

The tools can be categorized as follows:

- **Diagnostic tools:** These tools provide functionality to perform diagnostic operation for data assimilation applications. For example, the efficiency of assimilation techniques can be assessed and statistics significance tests can be performed.
- **Perturbation tools:** These tools allow the user to generate perturbations with prescribed properties in order to generate ensembles of model states or to perform perturbed ensemble integrations.
- **Transformation tools:** Assimilation algorithms that base on the Kalman filter assume Gaussian distributions for optimality. These tools provide functionality to perform preliminary transformations of variables or ensembles to improve the performance with non-Gaussian distributions.

- **Utilities:** Tools of this category provide additional functionality. Examples are tools for efficient manipulations of sparse matrices or the treatment of particular observation types.
- **Analysis steps:** Tools to compute the analysis step of an ensemble data assimilation method.

The release V0 of the SANGOMA tools represented a collection of tools that exist among the partners of SANGOMA. They were not yet adapted to the final standard interface structure, which is documented in the Deliverable D1.5. The following release V1 included tools that are adapted to the standard data model and interface structure. Both tools from the V0 release as well as additional tools were included. All tools were independent of the particular data-assimilation framework and should be usable with any other implementation of data-assimilation algorithms. Next to the documentation in this document, the tool release V1 also included examples on how to use a particular tool.

The final release V2 of the SANGOMA tools extended the collection of tools. The implementation work focused on diagnostic tools. However, it was also decided to add the analysis step of several variants of ensemble Kalman filters. With this, a user can also compute an analysis step using some ensemble of model states and a set of observation. These analysis steps are simplified compared to those implemented in the different available tools boxes. For example, they don't provide parallelization and dynamic allocation of internal arrays. With these simplifications, the tools are usable for moderately-sized problems (e.g. state dimensions of  $O(10^5)$ ), but for larger systems and for higher efficiency, we recommend to use one of the data assimilation toolboxes that have been described in DL1.1 of SANGOMA.

The tools release package can be downloaded from the project web site at <http://www.data-assimilation.net/Tools/>. In addition, the collection of tools is available at Source Forge at <http://sourceforge.net/projects/sangoma/> where it will remain available after the end of the SANGOMA project.

### 2.2.1 Overview of tools

#### Diagnostic Tools

##### *Fortran*

|                                      |  |
|--------------------------------------|--|
| <b>sangoma_CheckEnsSpread</b>        | Compute ensemble spread and deviation of ensemble mean from an input state |
| <b>sangoma_CheckNormality</b>        | Anderson-Darling Test to check normality of a sample                       |
| <b>sangoma_CheckWhiteness</b>        | Check whiteness of innovations   |
| <b>sangoma_CompareObsDiag</b>        | Compare observation-space diagnostics                                      |
| <b>sangoma_ComputeBRIER</b>          | Compute the Brier skill score and its decomposition, and the entropy       |
| <b>sangoma_ComputeCRIGN</b>          | Compute CRPF and CRIGN scores  |
| <b>sangoma_ComputeCRPS</b>           | Compute the CRPS and its decomposition                                     |
| <b>sangoma_ComputeEffSample</b>      | Compute the effective sample size of a particle filter                     |
| <b>sangoma_ComputeEnsStats</b>       | Compute ensemble statistics  |
| <b>sangoma_ComputeHistogram</b>      | Compute ensemble rank histograms   |
| <b>sangoma_ComputeInvStats</b>       | Compute innovation statistics  |
| <b>sangoma_ComputeMutInf</b>         | Compute the mutual information   |
| <b>sangoma_ComputeRCRV</b>           | Compute the bias & the dispersion of the RCRV                              |
| <b>sangoma_ComputeRE</b>             | Calculate the relative entropy   |
| <b>sangoma_ComputeSMatrix</b>        | Compute scaled ensemble observation anomalies                              |
| <b>sangoma_ComputeSensitivity</b>    | Calculate the sensitivity matrix with <b>H</b> as matrix                   |
| <b>sangoma_ComputeSensitivity_op</b> | Calculate the sensitivity matrix with <b>H</b> as operator                 |
| <b>sangoma_arm</b>                   | Calculate array modes  |
| <b>sangoma_armca</b>                 | Check the consistency of an ensemble using array modes                     |
| <b>sangoma_ObsDiag</b>               | Compute sampled observation-space diagnostics                              |



**Matlab/Octave**

|                           |  |
|---------------------------|--|
| <b>computeBRIER</b>       | Compute the Brier skill score and its decomposition, and the entropy       |
| <b>computeCRPS</b>        | Compute the CRPS and its decomposition                                     |
| <b>computeRCRV</b>        | Compute the bias & the dispersion of the RCRV                              |
| <b>computeHistogram</b>   | Compute ensemble rank histograms   |
| <b>mutual_information</b> | Compute mutual information in a particle filter                            |
| <b>relative_entropy</b>   | Compute relative entropy in a particle filter                              |
| <b>sensitivity</b>        | Compute sensitivity of posterior mean to observations in a particle filter |

**Perturbation Tools**

**Fortran**

|                            |  |
|----------------------------|--|
| <b>sangoma_pseudornd2D</b> | Generate random fields with given correlation length |
| <b>sangoma_MVNormalize</b> | Perform multivariate normalization                   |
| <b>sangoma_EOFcovar</b>    | Initialize covariance matrix from EOF decomposition  |

**Matlab/Octave**

|  |  |
|--|--|
| <b>Weakly constrained ensemble perturbations</b> | Create ensemble perturbations that have to satisfy an a priori linear constraint |
|--|--|

**Transformation Tools**

**Fortran**

|                                 |   |
|---------------------------------|---|
| <b>sangoma_Anamorphosis</b>     | Computes local Gaussian anamorphosis                  |
| <b>sangoma_ComputeQuantiles</b> | Computes ensemble quantiles as input for anamorphosis |

**Matlab/Octave**

|  |  |
|--|--|
| <b>Empirical Gaussian Anamorphosis</b> | Determine the empirical transformation function such that a transformed variable follows a Gaussian distribution |
|--|--|

## Utilities

### *Fortran*

|                           |   |
|---------------------------|---|
| <b>sangoma_computepod</b> | Computes dominant POD modes from an ensemble of snapshots                             |
| <b>sangoma_costgrad</b>   | Computes the values of Objective function and Gradient using reduced state dimensions |
| <b>mod_sangoma_utils</b>  | Module of utilities for easy porting from Matlab                                      |

### *Matlab/Octave*

|                         |  |
|-------------------------|--|
| <b>hfradar_extractf</b> | Observation operator for HF radar surface currents |
|-------------------------|--|

## Analysis

### *Fortran*

|  |   |
|--|---|
| <b>sangoma_ens_analysis</b>            | Computes the analysis ensemble using the ETKF scheme  |
| <b>sangoma_local_ensemble_analysis</b> | Computes the local analysis ensemble using the ETKF scheme  |
| <b>sangoma_enkf_analysis</b>           | Compute analysis ensemble using the EnKF with perturbed observations (globally or with covariance localization)         |
| <b>sangoma_ensrf_analysis</b>          | Compute analysis ensemble using the EnSRF with serial observation processing (globally or with covariance localization) |
| <b>sangoma_estkf_analysis</b>          | Compute analysis ensemble using the global ESTKF method   |
| <b>sangoma_etkf_analysis</b>           | Compute analysis ensemble using the global ETKF method  |
| <b>sangoma_letkf_analysis</b>          | Compute analysis ensemble using the ESTKF method with observation localization  |
| <b>sangoma_letkf_analysis</b>          | Compute analysis ensemble using the ETKF method with observation localization   |
| <b>sangoma_netf_analysis</b>           | Compute analysis ensemble using the NETF method   |

**Matlab/Octave**

|  |   |
|--|---|
| <b>sangoma_ensemble_analysis</b>       | Computes the analysis ensemble using the EnSRF, EAKF, ETKF, ETKF2, SEIK, ESTKF or EnKF scheme                             |
| <b>sangoma_local_ensemble_analysis</b> | Computes the local analysis ensemble using the EnSRF, EAKF, ETKF, ETKF2, SEIK, ESTKF or EnKF scheme (domain localization) |
| <b>sangoma_local_EnKF</b>              | Computes the local analysis ensemble using the EnKF (covariance localization)   |

**2.2.2 Using the tools**

**Directory structure of the release**

In the code package of the release we distinguish in between tools implemented in Fortran and tools for use with Matlab or Octave. The directory structure is as follows:

```

Fortran/
  ---- diagnostics/
                ----- examples/
  ---- perturbations/
                ----- examples/
  ---- transformations/
                ----- examples/
  ---- utilities/
                ----- examples/
  ---- analysis/
                ----- examples/
Matlab/
  ---- diagnostics/
                ----- examples/
  ---- perturbations/
                ----- examples/
  ---- transformations/
                ----- examples/
  ---- utilities/
                ----- examples/
  ---- analysis/
                ----- examples/
    
```

The directories are named after the four categories. For each category there is a sub-directory `examples/` in which example implementations are included that show how to use a particular tool.

All tools come with a complete description and examples.

The main deliverable of WP2 is DL2.5 presenting a detailed technical description of the tools. For the mathematical background of the methods, WP3 provides the necessary information.

### 2.2.3 Deviations from DOW

None

#### Highlight:

The collection contains a nice mix of advanced and simple routines, including standard EnKF versions. Of particular interest to users already exploiting an assimilation tool is the collection of advanced diagnostic tools and the anamorphosis transformation tool. DL2.5 provides detailed information.

## 2.3 WP3: Innovative Data Assimilation techniques

This WP has been focusing on stochastic DA methods, developments of new DA methods (including non-linear observation operators), comparisons of non-Gaussian assumptions and developing algorithms for assessing observing systems. This WP had the most academic freedom of research leading to interesting discussions within the project and new lines of research.

Each group made significant progress in advanced data-assimilation techniques which will pave the way for additional features in future operational data assimilation.

### 2.3.1 Scientific and technical highlights

Classical particle filter methods are subjected to the curse of dimensionality, demanding the ensemble size to grow exponentially with the problem dimension size. A particle filter method trying to avoid the curse of dimensionality was thus developed further: The equivalent weight particle filter (EWPF) tries to avoid the collapsing of the ensembles by keeping a relatively small number of particles with equal weights at analysis time. For this, simple nudging terms are added to the equation to provide a proposal density, which modify the particle weights, and a resampling technique exploiting the maximum weight each particle can achieve. The calculation of this maximum weight is possible analytically only if the observing operator is linear. The method was successfully applied to a 2D case with a nonlinear barotropic flow. For cases with nonlinear observing operators, an iterative solver to find the maximum weights was implemented

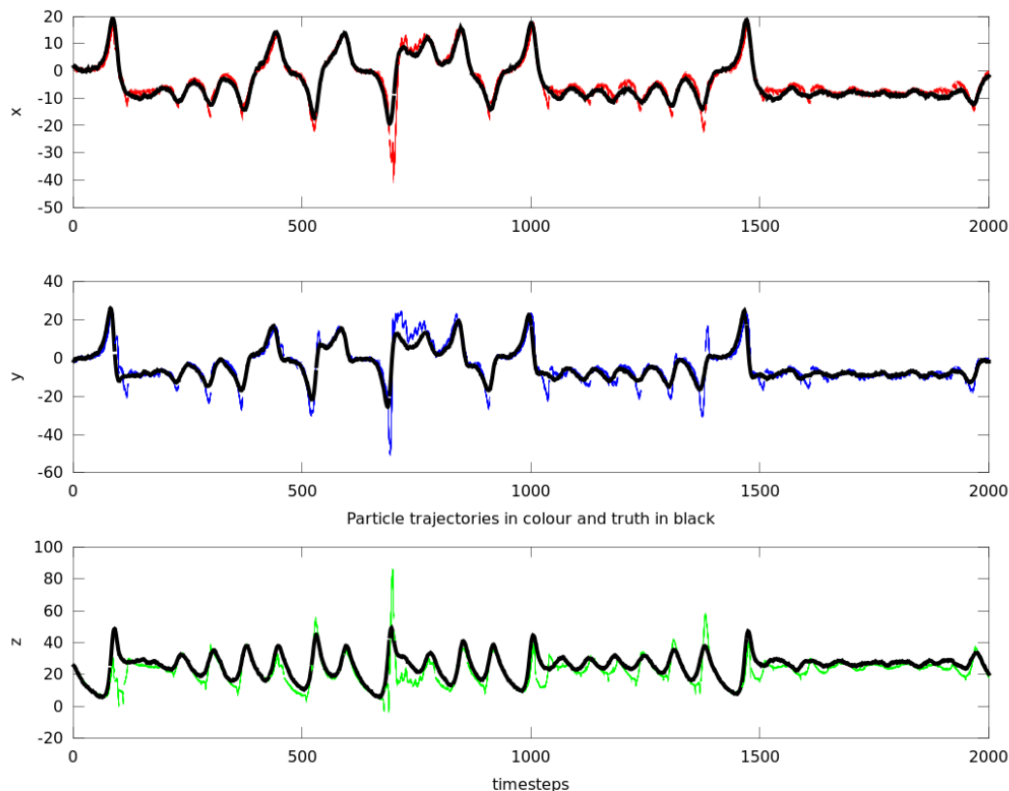


Figure 2.1: Results of EWPF application to the Lorenz 63 system when observing  $yz$ .

and tested on the chaotic Lorenz 63 system. The most challenging non-linear observations in this case are those which can lead to two different states (of different signs) because the observation of the product for example does not allow to distinguish which combination to retain. For details see <http://www.data-assimilation.net/Events/Year3/WP3.pdf> (Presentation at progress meeting) and <http://www.data-assimilation.net/Events/Year3/SANGOMASanita.pdf> (Presentation on nonlinear H)

A Multivariate Rank Histogram Filter (MRHF) exploits joint probability density functions to infer unobserved variables from observed ones. When working in a low dimensional system, the approach performs very well for non-gaussian problems, but in higher dimensions some strong hypotheses need to be formulated. By neglecting unobserved variables in some of the conditional statements a feasible MRHF is within reach but good insight in the effect of the hypotheses is needed. For details see <http://www.data-assimilation.net/Events/Year3/MRHF.pdf> (Presentation at progress meeting) and [http://das6.umd.edu/program/Daily/slides/2.2-Metref\\_Sammy.pdf](http://das6.umd.edu/program/Daily/slides/2.2-Metref_Sammy.pdf) (Poster)

An original method to include stochastic perturbations into a model was

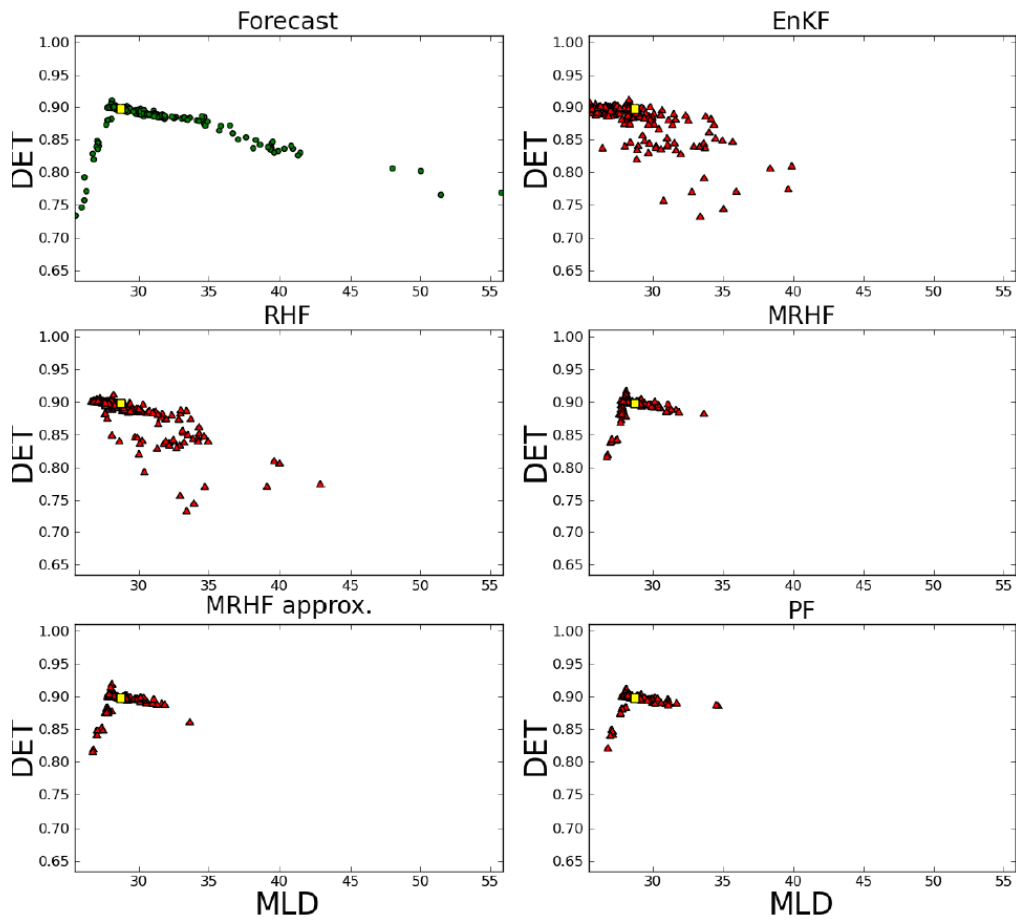


Figure 2.2: Application of the MRHF to the NATL025 implementation with biochemical model. Chl is observed and mixed layer depth and detritus are analysed. The members of the MRHF with or without the approximation in the joint probabilities work very well in this case.

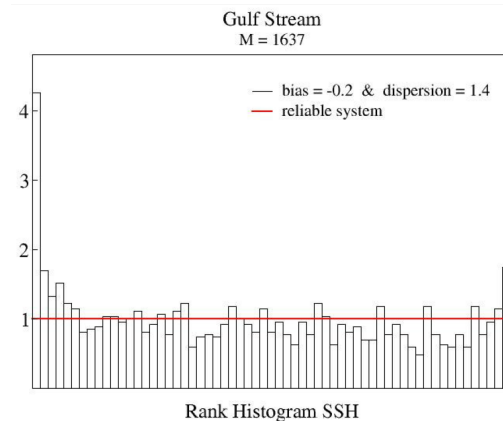


Figure 2.3: Application of stochastic perturbations on the state equation of the NATL025 implementation. SSH is observed and the rank histogram exhibits the nice ensemble spread.

developed. By using stochastic perturbations applied to the state equation one retrieves perturbations in which heat and salt are conserved. Also the perturbations lead to a nice and stable ensemble spread. For details see <http://www.data-assimilation.net/Events/Year3/brankartGC.pdf> (Presentation at progress meeting) and <http://dx.doi.org/10.1016/j.ocemod.2013.02.004> Brankart (2013).

Ensemble methods heavily exploit the reduced rank of the covariance matrix for efficient matrix inversions via the famous Woodbury formula. A new method combining such reduced rank matrices with local parametric covariances was developed to be able to combine large scale processes well modeled ensemble covariances the with smaller scale processes generally well modeled by parametric covariance functions. The combination normally does not lend itself to efficient inversion but with an iterative approach it still leads to efficient computations. In addition it allows for scale separations in the analysis. For details see <http://www.ocean-sci-discuss.net/11/895/2014/osd-11-895-2014.pdf> Beckers *et al.* (2014b).

Anamorphosis (data based change of variables) is a way to deal with data having a non-gaussian distribution, by transforming the data such that the transformed variables have a gaussian distribution. The transformation was tested in a coupled physical-biochemical model and indicates better spatial correlations when using the transformation. Therefore better corrections during assimilations can be expected. For details see <http://www.ocean-sci.net/8/121/2012/os-8-121-2012.pdf> Brankart *et al.* (2012).

For operational systems, error calculations can be quite time consuming when Optimal interpolation or 3DVar approaches are used. Simplified and approximate methods can alleviate the problem as shown in Beckers *et al.* (2014a).

The main deliverables of the workpackage are a state of the art (DL3.1)

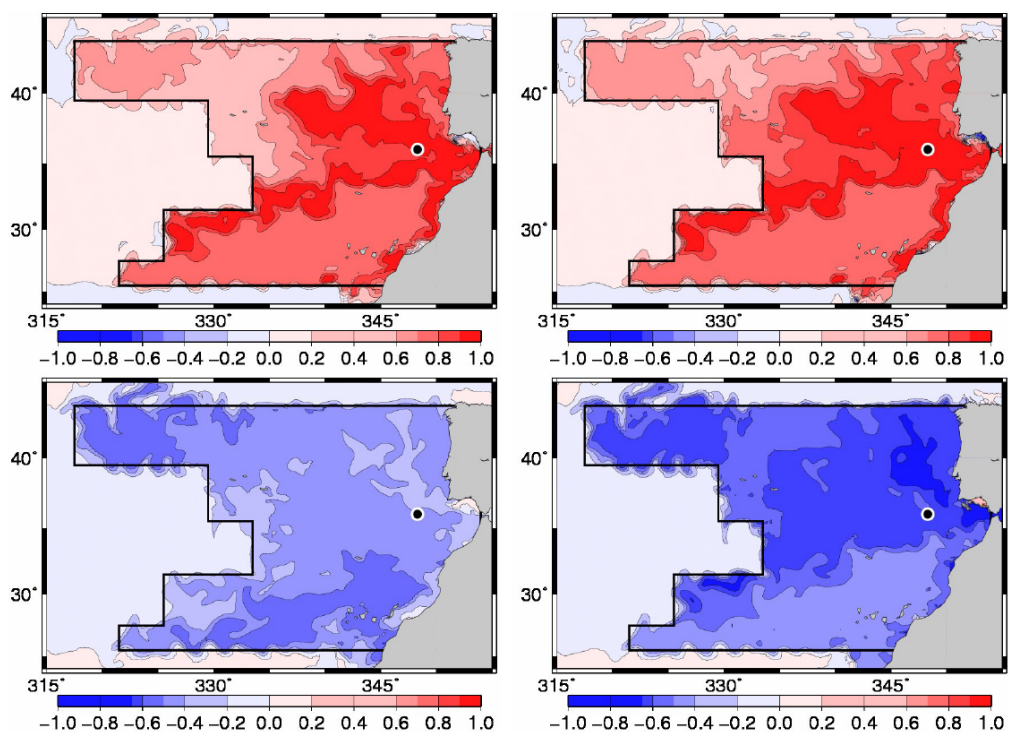


Figure 2.4: Phytoplankton (top panels) and nitrate (bottom panels) horizontal correlation structure with respect to phytoplankton without anamorphosis (left panels), and after local anamorphosis transformations (right panels).



described in a coherent framework (link between different EnKF version), the EWPF implementation in several toolboxes used as a common implementation of an advanced method (DL3.3) and new methods at exploratory level (DL3.4) or included in the SANGOMA collection (DL3.5).

### 2.3.2 Deviations from DOW

None

#### Highlight:

Working version of EWPF in several toolboxes using the defined standards (proof of concept of both the DA method and the SANGOMA data model/interface), uniform description of state of the art DA methods and several new advanced DA techniques.

## 2.4 WP4: Benchmarks

This workpackage aimed to provide some well documented and representative test cases for comparing data assimilation techniques. It also focused on adequate metrics to assess quality in non-linear regimes. The first year was concentrating in defining the details of the setup of the benchmark models. As planned, a small size (Lorenz 96), medium size (double gyre) and large size problem (Atlantic Ocean) was formulated. The setup of the benchmark models is completely specified and documented and some basic metrics are defined. A list of metrics not relying on the Gaussian-distribution assumption have also been prepared and discussed within the group and will serve as a basis for future assessment. It is also worth noting that the benchmark setup and use within an assimilation toolbox seems quite well prepared, as internal tests showed that a newcomer (both for NEMO and the toolbox) could reach a fully working ensemble assimilation within three month.

All benchmarks have been fully documented and implemented and executed by at least two partners. Details on which perturbations and ensemble members to generate were discussed and specified.

### 2.4.1 Scientific and technical highlights

In the first reporting period, the 3 SANGOMA benchmarks (small, medium and large case) had been defined, together with statistical metrics to evaluate stochastic assimilation methods (DL4.1).

Then, these 3 benchmarks have been implemented by the SANGOMA partners to evaluate various kinds of assimilation methods.

- Deliverable 4.2 lists the benchmarks that are being implemented by every SANGOMA partner, and provides the list of assimilation methods that will be evaluated by each partner.
- Deliverable 4.3 provides detailed information on how the probabilistic metrics are applied on small and medium benchmarks.

At the SANGOMA progress meeting of April 2014, presentations of the various benchmark implementations have shown that valuable results have already been obtained. In particular :

- Five implementations of the medium case benchmark allowed to intercompare different assimilation methods.
- Two implementations of the large scale benchmark (by GHER and LGGE as planned in the project) have been tested with a 6-month assimilation experiment (assimilating real-world ocean observations). Evaluation of these experiments has been performed, by applying more physical metrics at GHER and more probabilistic metrics at LGGE. Also the two implementations use complementary perturbations (density or external forcing).

In addition to standard metrics used to assess the quality of forecasts, the following probabilistic metrics haven been used:

- Rank Histogram
- Reduced Centered Random Variable (RCRV)
- Continuous Ranked Probability Scores (CRPS)
- Brier score Entropy

It has been shown that they provide additional insight into the quality of the simulations and therefore have been added to the SANGOMA tools list.

#### 2.4.2 Deviations from DOW

None

#### Highlight:

Fully documented benchmarks, also taken up by NEMOVAR, a series of probabilistic metrics documented and included into the toolbox. Main deliverable: Benchmark description (DL4.1) and use of new metrics on benchmarks, including probabilistic ones. Among the highlights we can also mention that the medium size benchmark was successfully implemented by a post-doc new to the assimilation toolbox and the NEMO model within 3 month.

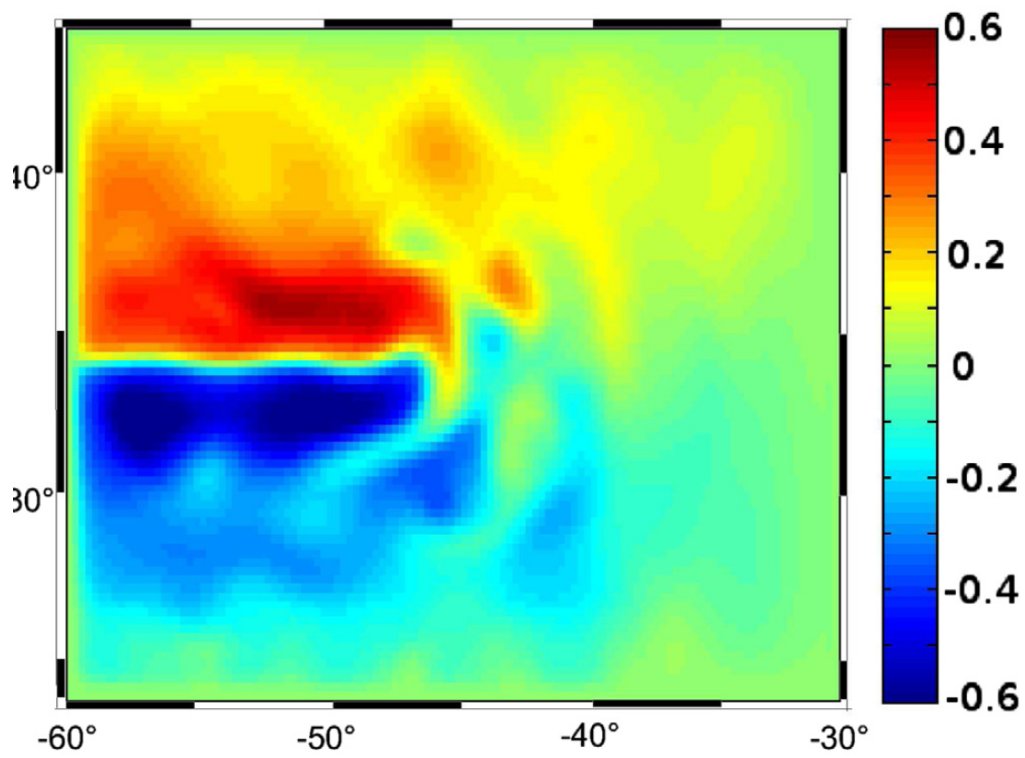


Figure 2.5: Snapshot of SSH in the medium-size benchmark.

## 2.5 WP5: Data Assessment

This WP tried to assess the impact of new remote sensed ocean data on the model state estimations and their potential in a data assimilation setup in a future operational context.

The new data types have been identified and characterized. Several model implementations have been achieved for assessing observing systems and impact of new data types.

### 2.5.1 Scientific and technical highlights

- Task 5.1 Identify new data types: This task has been completed by the Delivery of DL5.1, a list of remote-sensed variables with their associated characteristics.
- Task 5.2 Assessing observing systems: The work from G. Candille, CNRS-LEGI evaluated the relative impact of different altimeters constellations (Envisat and/or Jason 1) for the reduction of uncertainties and the improvement of model skills for unobserved temperature and salinity profiles. The work has used the large-scale NEMO benchmark from WP4. DL5.2 reported on the impact of new ecosystem data, although the contents is more related to ocean physics than ecosystem, consistently with the work performed in WP4.
- Task 5.3 Large-scale models: This task has build up on the synthetic data experiments carried out by G. Candille (CNRS-LEGI) in WP4. The results on assimilating real along-track altimeter data (Altika and SWOT) were analyzed in DL5.3 and DL5.4 with large-scale ocean models. In particular correlated observational errors were analyzed and resulted in a way to transform data so that a diagonal covariance matrix can be used.
- Task 5.4 Regional scale models: The work used a 1/60th deg ROMS model of the Ligurian Sea, using realistic forcings and real observations from two WERA HF radar systems operating during 2009 and 2010. The initial work has focused on the representation of model errors and covariance localization using as success criteria both the statistical robustness of increments and the expected error reduction. Localization is often applied for ensemble methods to avoid long range corrections introduced by artificial correlations found when under-sampling due to the number of members of the ensemble. Two problems which occur in this context can now be tackled: how to build the localization function objectively and how to maintain global conservation constraints when adding localization. For the limitation in space of the corrections based on an observation, the criterium is that the expected improvement brought by the correction should be larger than the error introduced by an incorrect specification of covariances. The latter effect is measured by a bootstrap method so that

Observations in the interior of the model domain

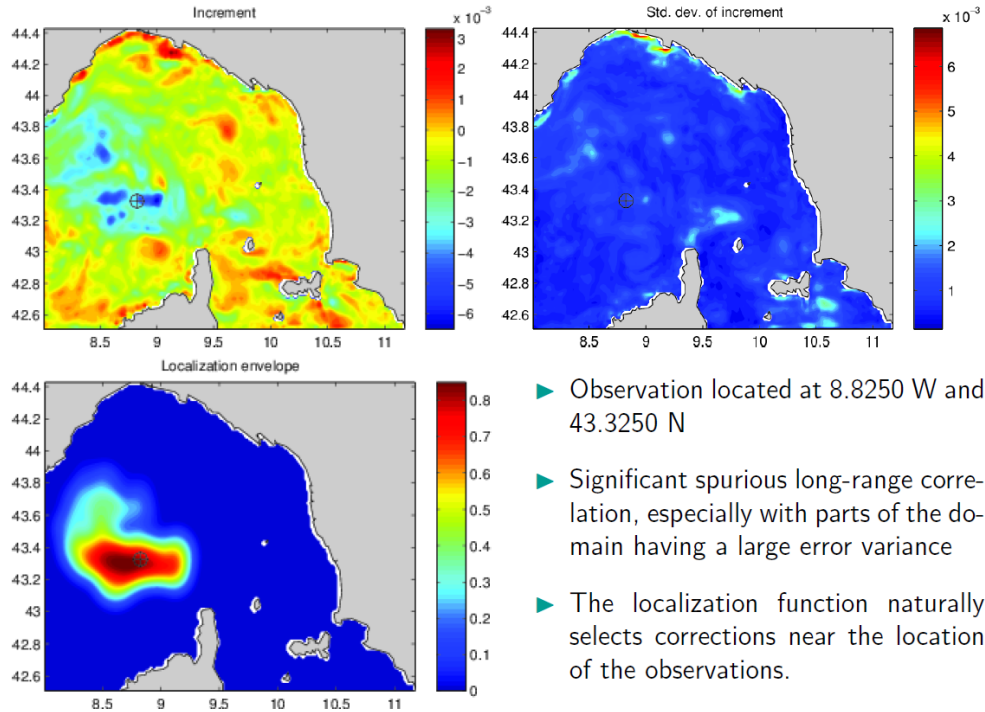


Figure 2.6: Application of the objective localisation determination in the ROMS Ligurian sea implementation. Velocity is observed by radar and the original increment shows spurious long range correlations particularly in regions where the error covariance is uncertain. With the objective localisation detection increments will be applied only in relevant points.

the localization function can be calculated. For maintaining conservation properties when applying the localisation approach, one has to modify the ensemble covariance such that the error variance on the conserved property is zero. This can be formulated mathematically rather easily for the forecast error but requires special care when updating the analysis error covariance. For details see <http://www.data-assimilation.net/Events/Year3/assimlocens.pdf> (Presentation at progress meeting) DL5.6 summarizes the results of a data assimilation experiment with a regional-scale ocean model in particular the duration of the correction’s impact and the different assimilation strategies (EnKS, vs AEnKF, window size).

Task 5.5 Lagrangian sea ice parameters The work has been initiated in an external collaboration with F. Massonnet at Université Catholique de Louvain la Neuve. Assimilation of satellite sea ice drift data in a NEMO-LIM model has been performed with the EnKF and global pa-

## Effect on ice drift velocities (Massonnet et al. in review)

**Partial success:  
Improved match to  
the observations  
assimilated.**

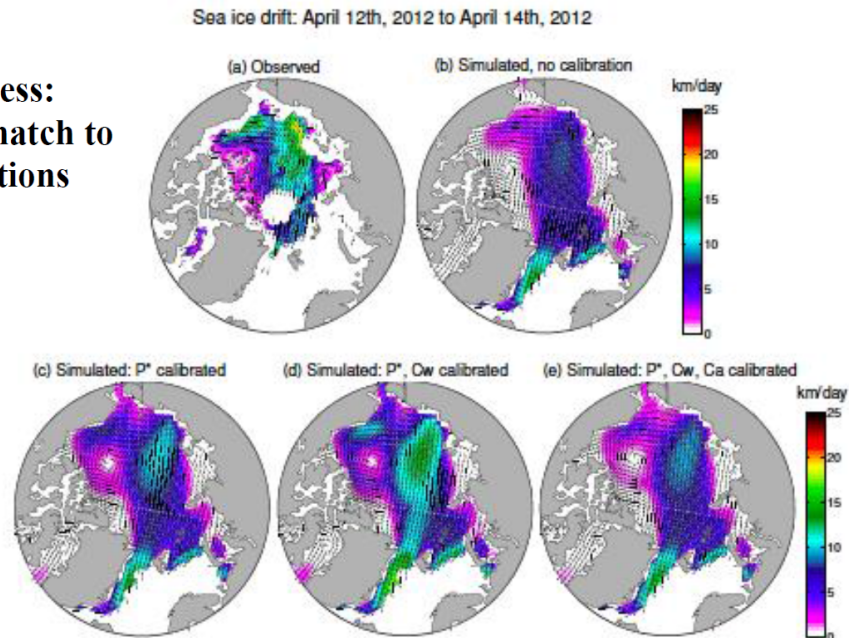


Figure 2.7: Assimilation of ice drift

Parameters of sea ice drag coefficients and sea ice strength (known as  $P^*$  in the (Elastic)-Viscous-Plastic sea ice rheology) have been estimated by a technique of state augmentation. The results show strong improvements for the estimation of two parameters (one drag coefficient and  $P^*$ ) but the only the ratio of air and ocean drag coefficient can be estimated correctly. DL5.7 discusses the results of the data assimilation experiment aiming to estimate Lagrangian sea ice parameters in a twin experiment scenario and using real observations. Estimating ice-related dynamic parameters with the assimilation tool reduced the errors in large areas.

**Task 5.6** Prior errors detection by observational arrays A diagnostic code has been tested by CNRS-LEGOS on a coastal oceanographic application and shared with CNRS-LEGI. DL5.8 provided the RMSpectrum library and results of array performance analyses.

Selecting among different locations and combinations of observing systems is simplified by characterizing the incremental information via the spectrum of the observed part of the error covariance of the forecast. This leads to selection criteria which allow to distinguish the added value of ferryboxes vs gliders for example

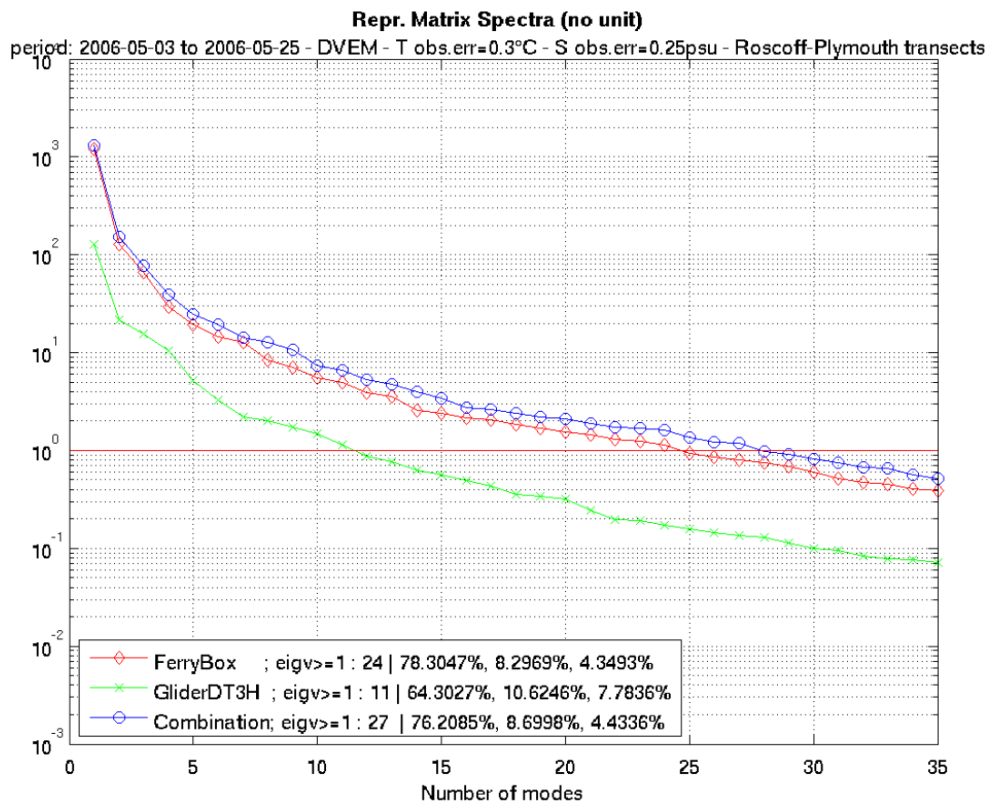


Figure 2.8: Eigenvalues of the representers from Ferrybox (red) remain larger than baseline from observations (flat line) for a larger number of eigenvalues than the one deducted for gliders (green line). Hence they potentially have a larger range of detectable modes

## 2.5.2 Deviations from DOW

None

### Highlight:

Description of new data types (DL5.1), general assessment tools (DL5.8), implementation of assimilation of new data types with interesting results:

- Along-track altimetry data (Envisat and Altika): stochastic parameterizations of the model dynamical uncertainties proved to be suitable
- Wide-swath sea surface height measurements (SWOT mission): Efficient method to account for the correlation of observational error
- HF Radar current observation: Space-time covariance suitable to capture inertial oscillation
- Sea-ice thickness: Useful to calibrate relevant parameters of the sea-ice

## 2.6 WP6: Knowledge transfer

This WP had to ensure that knowledge gained from the project and tools developed in it will be exploited outside of the SANGOMA consortium. Among the general activities we find the releases the code, publication of papers and presentation of the project partners at different meetings.

An important point is that the tools developed by SANGOMA are useful to the scientific and operational community:

- For scientific groups, the open sharing of tools and benchmarks will allow other scientist to test their new ideas and their effect by comparing them to results from SANGOMA. Also reports on DA techniques and new data types will speed up uptake of new techniques by scientist using data assimilation. We reached this community by standard scientific publications and communications but also via the Ph.D workshops, schools and student exchanges.
- myOcean: a strong involvement of myOcean representatives in SANGOMA was planned and implemented (two partners are



also partners in myOcean, one as MFC; the steering committee of SANGOMA includes E. Dombrowsky as representative of myOcean; already two of the variable advisory-board members emerged from myOcean: D. Obaton and C.E. Testut; a specific consortium agreement between SANGOMA and myOcean was set up; myOcean science days workshop on data assimilation was organised with SANGOMA.)

- Other GMES service evolution projects showed interest (OSS2015 and myWave for data assimilation techniques) but no particular action has emerged.
- Operational centers (within myOcean and outside) should find in the SANGOMA toolboxes and benchmarks easy ways to test new techniques before deciding to optimise them for their operational implementation. Diagnostic tools should also be of direct interest. To make sure our work is relevant we submitted a list of tools to be developed to myOcean.
- Private firms have shown their interest in using tested tools as they generally have less time for reimplementing documented tools. For these firms, we kept them informed and invited them to the Liège Colloquium.
- Contacts with ESA (ESFRI) through the advisory board participation at the kick-off meeting were kept alive, notably to prepare the ESA summerschool contribution from SANGOMA through the UREAD group. NERSC is also involved in ESA's Ocean Colour CCI in which he uses data assimilation techniques further developed within SANGOMA.
- a workshop with operational users was planned at month 12. During GMES marine projects coordination meeting on 24/05/2012 in Brussels we discussed this matter with the myOcean group and it was agreed that we should reach this community together during the myOcean Science Days. Therefore, invitations were send via the myOcean contact list but also directly from the SANGOMA side to a dozen of additional institutes interested in operational modelling (like Actimar and MUMM).
- During the myOcean Science Days, 60-70 participants attended the workshop on data assimilation. Four oral presentations and four posters showed the SANGOMA approach. A call for feedback was also launched and a survey <http://www.surveymonkey.com/s/ZX3P9D8> indicating very positive views on our proposed data model and interfaces.
- The Liège colloquium 2015 focused on DA and SANGOMA took a leading role there, gathering almost 200 scientists.
- Web pages of SANGOMA have been updated regularly (see DL6.2) and include scientific highlights.

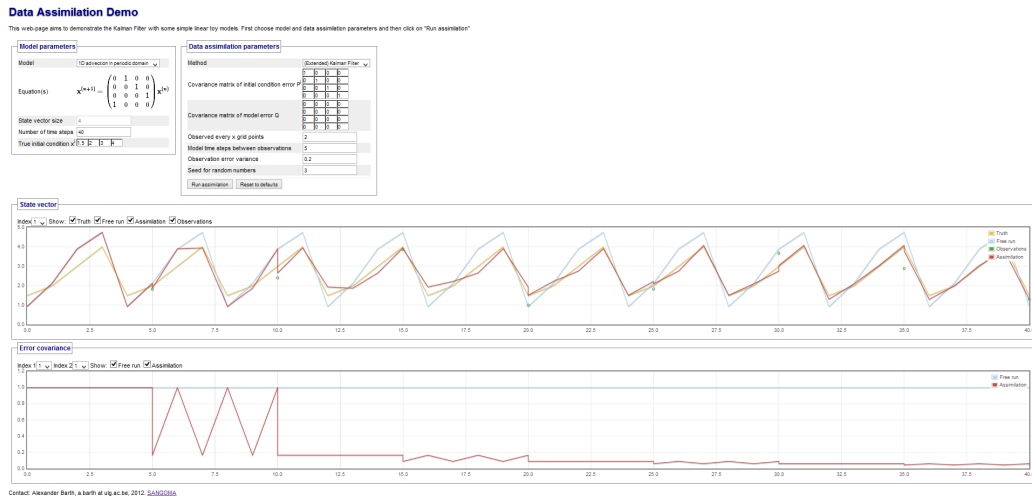
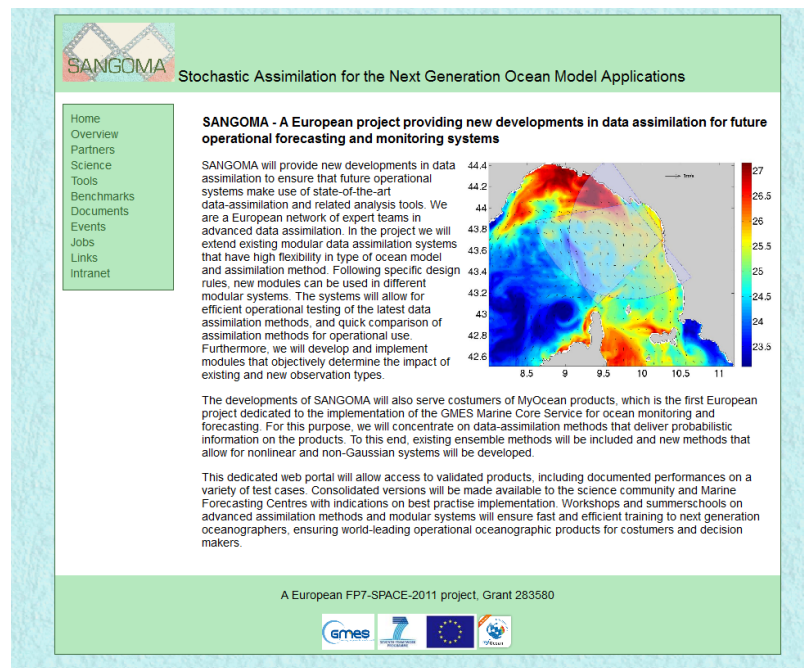


Figure 2.9: Web interface to data assimilation demonstration

- For demonstration purposes, a web interface to some assimilation techniques was installed to illustrate the concepts: <http://www.data-assimilation.net/Tools/AssimDemo/>



SANGOMA web site <http://www.data-assimilation.net>.

### 2.6.1 Deviations from DOW

Less workshops than planned were organized, due to difficulty to reach and mobilize stakeholders to join.

**Highlight:**

Software distributed (500 downloads at the project end; the bundled V2 software was also released with the documentation DL6.14), SVN server still up and being updated, 87 different dissemination activities, Liège Colloquium 2015, web based DA demonstration tool.

## 2.7 WP7: Management

The only significant deviation from the plan was the delay in final report delivery.

## Chapter 3

# Evolutions

A problem posed by the end of the SANGOMA project is the durability of our collaborative approach.

The scientists involved will certainly continue to collaborate but the future of the SANGOMA software and standards needs attention.

On a short term, ULg will maintain the SVN servers and web pages and in view of the permanent positions of the scientists at ULg involved, the maintenance should be easily ensured during the next decade.

For adding content ULg will be able to make small bug corrections but for major enhancement new projects are needed. A few possible ways to prepare such projects are the following

- The report of the 4th NEMO-ASSIM meeting January 29th 2015, included reference to the SANGOMA approach and also the desire to come up with a reference setup for testing DA with NEMO, very close to our medium size benchmark.
- Partners should propose their (national or other) projects highlighting they will feeding back their new findings into the SANGOMA toolbox: The benefit for the proposal is to show its uptake strategy and the benefit for the SANGOMA consortium is the sustainability of the consortium. This approach was already taken in the past with an OSTST-ASSIM for the French groups EUMETSAT/CNES call.
- Sangoma WP5 links to the Sentinel missions should be exploited: Task 5.3 prepares for using future ocean altimeter onboard the Sentinel-3 mission. Task 5.5 prepares for using sea ice altimeter onboard the Sentinel-3 mission.
- Arising opportunities to include the methods in Copernicus Marine Services in H2020.
- Ensure GMEMS participations.

- SMEs using the toolbox might wish to include some of their work into the toolbox (as it would provide additional testing from other scientists).


















New problems to be tackled in the future will include assimilation in unstructured grid models and coastal benchmarks, as well as the problem of big data assimilation.

## Chapter 4

## Other information

The publication list originating from the project at the end of it is provided from the EU portal. For updated version, with papers published later, please visit <http://www.data-assimilation.net/Documents/>. Dissemination activities are also those reported to EU.

Project Publications (Peer reviewed publication)

| Order No | D.O.I.  | Title  | Author(s)  | Title of the periodical or the series                 | Number, date or frequency | Publisher   | Place of publication | Date of publication | Relevant pages | Open access is/will be provided to this publication | Status    | Actions   |
|----------|---|--|--|---|---------------------------|---|----------------------|---------------------|----------------|---|-----------|---|
| 1        | <a href="https://doi.org/10.5194/oss-8-141-2012">10.5194/oss-8-141-2012</a>                 | Towards an improved description of ocean uncertainties: effect of local anamorphic transformations on spatial correlations   | J.-M. Brankart, C.-E. Testut, P.-A. Beau, M. Beaudouin, P. Brasseur, and J. Verron                   | Ocean Science   | Vol. 8, issue 2, 2012     | Copernicus Publications on behalf of the European Geosciences Union |                      | 06/03/2012          | 121-142        | Yes   | VALIDATED |    |
| 2        | <a href="https://doi.org/10.1016/j.osecomod.2013.11.002">10.1016/j.osecomod.2013.11.002</a> | Comparison of different assimilation schemes in a sequential Kalman filter assimilation system   | Y. Yan, A. Barth, J.M. Beckers   | Ocean Modelling                                       | Vol. 73                   | null  |                      | 01/01/2014          | 123-137        |   | VALIDATED |    |
| 3        | <a href="https://doi.org/10.1016/j.leageo.2012.03.026">10.1016/j.leageo.2012.03.026</a>     | Software for ensemble-based data assimilation systems—Implementation strategies and scalability  | L. Nerger and W. Hiller  | Computers and Geosciences                             | 55                        | Elsevier Limited  | Germany              | 01/06/2013          | 110-118        |   | VALIDATED |    |
| 4        | <a href="https://doi.org/10.5194/amt-8-11-895-2014">10.5194/amt-8-11-895-2014</a>           | Multi-scale optimal interpolation: application to DINEOF analysis speed with a local optimal interpolation   | J.-M. Beckers, A. Barth, I. Tomazic, A. Alvera-Azcárate  | Ocean Science Discussions                             | Vol. 11/Issue 2           | European Geosciences Union  | Germany              | 01/01/2014          | 895-941        |   | VALIDATED |    |
| 5        | <a href="https://doi.org/10.1175/JTECH-D-13-00130.1">10.1175/JTECH-D-13-00130.1</a>         | Approximate and Efficient Methods to Assess Error Fields in Spatial Gridding with Data Interpolating Variational Analysis (DIVA)   | Jean-Marie Beckers, Alexander Barth, Charles Troupin, Alda Alvera-Azcárate                           | Journal of Atmospheric and Oceanic Technology         | Vol. 31/Issue 2           | American Meteorological Society                                     | United States        | 01/02/2014          | 515-530        |   | VALIDATED |    |
| 6        | <a href="https://doi.org/10.1175/MWR-D-13-00246.1">10.1175/MWR-D-13-00246.1</a>             | On the Choice of an Optimal Localization Radius in Ensemble Kalman Filter Methods  | Paul Kirchgesner, Lars Nerger, Angelika Bunse-Gerstner   | Monthly Weather Review                                | Vol. 142/Issue 6          | American Meteorological Society                                     | United States        | 01/06/2014          | 2165-2175      |   | VALIDATED |    |
| 7        | <a href="https://doi.org/10.1016/j.osecomod.2013.02.004">10.1016/j.osecomod.2013.02.004</a> | Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modelling  | Jean-Michel Brankart   | Ocean Modelling                                       | Vol. 66                   | null  |                      | 01/06/2013          | 64-76          |   | VALIDATED |    |
| 8        | <a href="https://doi.org/10.1175/MWR-D-14-00392.1">10.1175/MWR-D-14-00392.1</a>             | Linking the Anomaly Initialization Approach to the Mapping Paradigm: A Proof-of-Concept Study  | Robin J. T. Weber, Alberto Carrasi, Francisco J. Doblas-Keyes  | Monthly Weather Review                                | Vol. 143/Issue 11         | American Meteorological Society                                     | United States        | 01/11/2015          | 4695-4713      |   | VALIDATED |    |
| 9        | <a href="https://doi.org/10.5194/amt-8-11-425-2015">10.5194/amt-8-11-425-2015</a>           | Assessment of an ensemble system that assimilates Jason-1/Envisat altimeter data in a probabilistic model of the North Atlantic ocean circulation                                      | G. Candille, J.-M. Brankart, P. Brasseur   | Ocean Science   | Vol. 11/Issue 3           | European Geosciences Union  | Germany              | 01/01/2015          | 425-438        |   | VALIDATED |    |
| 10       | <a href="https://doi.org/10.1002/2014JC010349">10.1002/2014JC010349</a>                     | Ensemble assimilation of ARGO temperature profile, sea surface temperature, and altimetric satellite data into an eddy permitting primitive equation model of the North Atlantic Ocean | Y. Yan, A. Barth, J. M. Beckers, G. Candille, J. M. Brankart, P. Brasseur                            | Journal of Geophysical Research: Oceans               | Vol. 120/Issue 7          | AGU   |                      | 01/07/2015          | 5134-5157      |   | VALIDATED |    |
| 11       | <a href="https://doi.org/10.5194/amt-8-21-869-2014">10.5194/amt-8-21-869-2014</a>           | A non-Gaussian analysis scheme using rank histograms for ensemble data assimilation  | S. Metref, E. Cosme, C. Snyder, P. Brasseur  | Nonlinear Processes in Geophysics                     | Vol. 21/Issue 4           | European Geosciences Union  |                      | 01/01/2014          | 869-885        |   | VALIDATED |    |
| 12       | <a href="https://doi.org/10.5194/amt-8-1285-2015">10.5194/amt-8-1285-2015</a>               | A generic approach to explicit simulation of uncertainty in the NEMO ocean model   | J.-M. Brankart, G. Candille, F. Garnier, C. Calonn, A. Melet, P.-A. Boubvier, P. Brasseur, J. Verron | Geoscientific Model Development                       | Vol. 8/Issue 5            | Copernicus GmbH (Copernicus Publications), EGU                      | Germany              | 01/01/2015          | 1285-1297      |   | VALIDATED |  |
| 13       | <a href="https://doi.org/10.1175/mwr-d-14-00375.1">10.1175/mwr-d-14-00375.1</a>             | Extending the Square Root Method to Account for Additive Forecast Noise in Ensemble Methods  | Patrick Mima Raanes, Alberto Carrasi, Laurent Berino   | Monthly Weather Review                                | Vol. 143/Issue 10         | American Meteorological Society                                     | United States        | 01/10/2015          | 3857-3873      |   | VALIDATED |  |
| 14       | <a href="https://doi.org/10.1002/2015JC009705">10.1002/2015JC009705</a>                     | Accounting for model error due to unresolved scales within ensemble Kalman filtering   | Lewis Mitchell, Alberto Carrasi  | Quarterly Journal of the Royal Meteorological Society | Vol. 141/Issue 689        | John Wiley and Sons Ltd   | United Kingdom       | 01/04/2015          | 1417-1428      |   | VALIDATED |  |
| 15       | <a href="https://doi.org/10.1002/2015JC009705">doi:10.1002/2015JC009705</a>                 | Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter  | F. Massonnet, H. Gosse, T. Fichefet, F. Couillon   | Journal of Geophysical Research: Oceans               | Vol. 119/Issue 7          | AGU   |                      | 01/07/2014          | 4168-4184      |   | VALIDATED |  |
| 16       |   | On the influence of model nonlinearity and localization on ensemble Kalman smoothing   | Lars Nerger, Svenja Schulte, Angelika Bunse-Gerstner   | Quarterly Journal of the Royal Meteorological Society | Vol. 140/Issue 684        | John Wiley and Sons Ltd   | United Kingdom       | 01/10/2014          | 2249-2259      |   | VALIDATED |  |
| 17       | <a href="https://doi.org/10.1175/MWR-D-15-0073.1">10.1175/MWR-D-15-0073.1</a>               | Assessment of a Nonlinear Ensemble Transform Filter for High-Dimensional Data Assimilation   | Julian Tödtter, Paul Kirchgesner, Lars Nerger, Bodo Ahrens   | Monthly Weather Review                                | Vol. 144/Issue 1          | American Meteorological Society                                     | United States        | 01/01/2016          | 409-427        |   | VALIDATED |  |





| No | Type of activities                      | Main leader   | Title   | Date       | Place  | Type of audience                                  | Size of audience | Countries addressed | Status    | Actions | Order |
|----|---|---|---|------------|--|---|------------------|---------------------|-----------|---------|-------|
| 1  | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | CNRS-LEGOS contribution to SANGOMA  | 24/10/2011 | Liège, Belgique (SANGOMA European Project-kick-off meeting)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 1     |
| 2  | Flyers                                  | UNIVERSITE DE LIEGE                                     | Stochastic Assimilation for the Next Generation Ocean Model applications  | 01/11/2011 | Belgium, United Kingdom, Germany, Netherlands, France, Norway  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 2     |
| 3  | Publication                             | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | Towards an improved description of ocean uncertainties: effect of local anamorphic transformations..  | 06/03/2012 | European Geosciences Union   | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 3     |
| 4  | Presentations                           | UNIVERSITE DE LIEGE                                     | Assimilation of high-frequency radar currents in the Ligurian Sea   | 18/04/2012 | Livorno, Italy (Workshop MOMAR)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 4     |
| 5  | Posters                                 | UNIVERSITE DE LIEGE                                     | Data assimilation in Ligurian Sea using HF radars   | 19/11/2012 | Geesthacht, Germany (MyOcean Science Days 2012)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 5     |
| 6  | Posters                                 | TECHNISCHE UNIVERSITEIT DELFT                           | The OpenDA data-assimilation toolbox  | 19/11/2012 | Geesthacht, Germany (MyOcean Science Days 2012)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 6     |
| 7  | Posters                                 | UNIVERSITE DE LIEGE                                     | Ensemble Data Assimilation in a global coupled sea ice model  | 19/11/2012 | Geesthacht, Germany (MyOcean Science Days 2012)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 7     |
| 8  | Posters                                 | THE UNIVERSITY OF READING                               | A fully nonlinear filter for high-dimensional systems   | 19/11/2012 | Geesthacht, Germany (MyOcean Science Days 2012)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 8     |
| 9  | Presentations                           | TECHNISCHE UNIVERSITEIT DELFT                           | Wave data-assimilation for SWAN using OpenDA  | 20/11/2012 | Geesthacht, Germany (MyOcean Science Days 2012)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 9     |
| 10 | Presentations                           | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | Advanced assimilation methods for the next generation of ocean monitoring and forecasting centres   | 20/11/2012 | Geesthacht, Germany (MyOcean Science Days 2012)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 10    |
| 11 | Presentations                           | ALFRED-WEGENER-INSTITUT FUER POLAR- UND MEERESFORSCHUNG | An ensemble-based forecasting system for the North and Baltic Seas using the BSH circulation model..  | 20/11/2012 | Geesthacht, Germany (MyOcean Science Days 2012)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 11    |
| 12 | Presentations                           | UNIVERSITE DE LIEGE                                     | Development of Stochastic Assimilation for the Next-Generation Ocean Model Applications   | 20/11/2012 | Geesthacht, Germany (MyOcean Science Days 2012)  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 12    |
| 13 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | Array testing and impact of observations in the coastal ocean by ensemble methods..   | 01/02/2013 | Lecco, Italy (Second International Workshop of the GODAE OceanView Coastal and Shelf Seas Task Team) | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 13    |
| 14 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | On non Gaussian Ensemble data assimilation  | 17/02/2013 | Banff, Probabilistic Approaches to Data Assimilation for Earth Systems                               | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 14    |
| 15 | Oral presentation to a scientific event | UNIVERSITE DE LIEGE                                     | Assimilation of simulated satellite altimetric data and ARGO temperature data into a double-gyre NEMO ocean model   | 07/04/2013 | EGU2013 Vienna   | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 15    |
| 16 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | Groupe de travail sur l'assimilation de données dans le Golfe de Gascogne et mers côtières  | 23/04/2013 | Toulouse, France (Journées 2013 du Groupe Mission MERCATOR-CORTOLIS)                                 | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 16    |
| 17 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | Impact of observations  | 24/06/2013 | Bologna, Italy (Bologna Advanced Data Assimilation Summer School)                                    | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 17    |
| 18 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | A non-Gaussian analysis scheme using rank histograms for ensemble data assimilation   | 07/10/2013 | 6th WMO SYMPOSIUM  | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 18    |
| 19 | Posters                                 | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | Array and model testing and impact of observations (with applications in the coastal ocean)   | 13/01/2014 | Toulouse, France (Colloque LEFE)   | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 19    |
| 20 | Posters                                 | ALFRED-WEGENER-INSTITUT FUER POLAR- UND MEERESFORSCHUNG | A comparison of linear and non-linear linear data assimilation techniques using a toy model   | 24/02/2014 | LMU Munich International Symposium on Data Assimilation 2014   | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 20    |
| 21 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE            | Stochastic array design with the sangoma_arm tool   | 01/04/2014 | Liège, Belgium (SANGOMA Project meeting)   | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 21    |
| 22 | Oral presentation to a scientific event | UNIVERSITE DE LIEGE                                     | Assimilation of ARGO temperature profiles, sea surface temperature and altimetric satellite data into an eddy permitting primitive equation model of the North Atlantic Ocean | 27/04/2014 | EGU 2014, Vienna   | Scientific community (higher education, Research) |                  | International       | VALIDATED |         | 22    |

| a scientific event                         | equation model of the North Atlantic Ocean   | Research  | International | VALIDATED | Research |
|--|--|---|---------------|-----------|----------|
| 23 Oral presentation to a scientific event | Advances in Climate Theory   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 24 Posters                                 | Assessment of stochastic filters for assimilation of high-frequency observations in a coupled physical-biological model of the Ligurian Sea  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 25 Posters                                 | The SANGOMA tools for data assimilation  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 26 Posters                                 | Extending NEMO For Ensemble Data Assimilation On Supercomputers With The Parallel Data Assimilation Framework PDAF   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 27 Posters                                 | Recent Progress In Nonlinear Data Assimilation   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 28 Posters                                 | Ensemble Assimilation Of ARGO Temperature Profile, Sea Surface Temperature And Altimetric Satellites Data Into An Eddy Permitting Primitive Equation Model Of The North Atlantic Ocean | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 29 Posters                                 | Towards data assimilation in a state-of-the-art physical-model of the North Atlantic: Estimation of model uncertainties using Stochastic parametrizations                              | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 30 Posters                                 | Reanalysis of the Southern Ocean with assimilation of sea surface temperatures, ice concentration and ice drift  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 31 Oral presentation to a scientific event | Modeling and data assimilation developments of the TOPAZ system in support of operational oceanography in the Arctic   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 32 Oral presentation to a scientific event | Ensemble-based array performance assessment  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 33 Oral presentation to a scientific event | EPSRC Centre for Doctoral Training - Mathematics of Planet Earth   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 34 Oral presentation to a scientific event | Ensemble-based array performance assessment in coastal ocean models  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 35 Oral presentation to a scientific event | Lessons learnt from 1200 EnKF cycles with TOPAZ.   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 36 Oral presentation to a scientific event | Extending the square root method to account for model noise in the ensemble Kalman filter  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 37 Oral presentation to a scientific event | Les Houches Ensemble data assimilation workshop  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 38 Oral presentation to a scientific event | Optimizing observation networks in the Bay of Biscay and English Channel   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 39 Posters                                 | Assimilation of sea surface temperature, sea ice concentration and sea ice drift in a model of the Southern Ocean  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 40 Oral presentation to a scientific event | Building Ensemble-Based Data Assimilation Systems for High-Dimensional Models  | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 41 Oral presentation to a scientific event | Non-Gaussian Data Assimilation Methods   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 42 Oral presentation to a scientific event | A comparison of linear and non-linear data assimilation methods using the NEMO ocean model   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 43 Oral presentation to a scientific event | Using EMPiRE to assess the impact of a fully non-linear data assimilation method with NEMO   | Scientific community (higher education, Research) | International | VALIDATED | Research |
| 44 Oral presentation to a scientific event | OpenDA-NEMO framework for Ocean Data Assimilation  | Scientific community (higher education, Research) | International | VALIDATED | Research |

|    |   |   |   |            |   |   |               |           |  |
|----|---|---|---|------------|---|---|---------------|-----------|--|
| 45 | Oral presentation to a scientific event | UNIVERSITE DE LIEGE                               | Local ensemble assimilation scheme with global constraints and conservation   | 07/05/2015 | University of Liège, Belgium (47th Liège Colloquium on Ocean Dynamics)        | Scientific community (higher education, Research) | International | VALIDATED |  |
| 46 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Data assimilation in a 4 <sup>th</sup> coupled physical-biochemical model of the North Atlantic using error modelling based on stochastic parameterizations of biogeochemical uncertainties | 07/05/2015 | University of Liège, Belgium (47th Liège Colloquium on Ocean Dynamics)        | Scientific community (higher education, Research) | International | VALIDATED |  |
| 47 | Oral presentation to a scientific event | UNIVERSITE DE LIEGE                               | Assimilation of HF radar in the Ligurian Sea: Spatial and Temporal scale considerations   | 07/05/2015 | Colloquium of Liège, Belgium (47th Liège Colloquium on Ocean Dynamics)        | Scientific community (higher education, Research) | International | VALIDATED |  |
| 48 | Oral presentation to a scientific event | UNIVERSITE DE LIEGE                               | Bias correction with data assimilation.   | 07/05/2015 | University of Liège, Belgium (47th Liège Colloquium on Ocean Dynamics)        | Scientific community (higher education, Research) | International | VALIDATED |  |
| 49 | Posters                                 | STIFTSELSEN MÅNSEN SENTER FOR MILDØOG FJERNMÅLING | A New Dual Ensemble Kalman Filter for State-Parameters Estimation in Subsurface Hydrology   | 09/05/2015 | International ENKF workshop, Flåm, Norway                                     | Scientific community (higher education, Research) | International | VALIDATED |  |
| 50 | Oral presentation to a scientific event | STIFTSELSEN MÅNSEN SENTER FOR MILDØOG FJERNMÅLING | Monitoring and Predicting Subsurface Organic Contaminants in the Port of Rotterdam using a Hybrid Ensemble Kalman Filter  | 09/05/2015 | International ENKF workshop, Flåm, Norway                                     | Scientific community (higher education, Research) | International | VALIDATED |  |
| 51 | Oral presentation to a scientific event | STIFTSELSEN MÅNSEN SENTER FOR MILDØOG FJERNMÅLING | Nonlinear Dynamics, Extremes, Geo-hazards and Predictability of the Earth System  | 26/05/2015 | University of Hamburg, GERMANY  | Scientific community (higher education, Research) | International | VALIDATED |  |
| 52 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Probabilistic Approaches (and Risk Assessment) in the Coastal Ocean   | 08/06/2015 | Biddeford, Maine, USA (Gordon Research Conference on Coastal Ocean Modelling) | Scientific community (higher education, Research) | International | VALIDATED |  |
| 53 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | AMICO 1a: Assimilation de données par méthode ensembliste et descente d'échelles dans le Golfe de Gascogne  | 15/06/2015 | Toulouse, France (Journées 2015 du GMMC)                                      | Scientific community (higher education, Research) | International | VALIDATED |  |
| 54 | Posters                                 | STIFTSELSEN MÅNSEN SENTER FOR MILDØOG FJERNMÅLING | Data assimilation by delay coordinate nudging   | 06/09/2015 | Dynamics Day EU   | Scientific community (higher education, Research) | International | VALIDATED |  |
| 55 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Data assimilation in a state-of-the-art physical-biochemical model of the North Atlantic: towards synergistic usage of Sea Level, SST and Ocean Colour observations                         | 13/05/2013 | University of Liège, Belgium (45th Liège Colloquium on Ocean Dynamics)        | Scientific community (higher education, Research) | International | VALIDATED |  |
| 56 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | The control of non-linear mesoscale ocean circulation through altimetric data assimilation revisited using a variational approach   | 07/10/2013 | Boulder, USA (2013 OSTST meeting)   | Scientific community (higher education, Research) | International | VALIDATED |  |
| 57 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Towards an ensemble strategy for altimetric data assimilation into eddy-resolving ocean circulation models  | 07/10/2013 | Boulder, USA (2013 OSTST meeting)   | Scientific community (higher education, Research) | International | VALIDATED |  |
| 58 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Towards an ensemble strategy for data assimilation into eddy-resolving ocean circulation models   | 04/11/2013 | Baltimore, USA (International GODAE OceanView symposium)                      | Scientific community (higher education, Research) | International | VALIDATED |  |
| 59 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Towards data assimilation in a coupled Physical-Biochemical model of the North Atlantic: Estimation of model uncertainties using stochastic parameterizations                               | 24/02/2014 | Honolulu, USA (AGU Ocean Science Meeting)                                     | Scientific community (higher education, Research) | International | VALIDATED |  |
| 60 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Ensemble data assimilation in a North Atlantic, eddy-permitting ocean circulation model using stochastic parameterization of the model dynamics   | 22/09/2014 | Toulouse, France (MyOcean Science Days 2014)                                  | Scientific community (higher education, Research) | International | VALIDATED |  |
| 61 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Stochastic estimation of parameters describing forcing uncertainties in a biogeochemical model  | 22/09/2014 | Toulouse, France (MyOcean Science Days 2014)                                  | Scientific community (higher education, Research) | International | VALIDATED |  |
| 62 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Ensemble assimilation of JASON/ENVISAT and JASON/AIRKA altimetric observations with stochastic parameterization of the model dynamical uncertainties  | 28/10/2014 | Lake Constance, Germany (2014 OSTST meeting)                                  | Scientific community (higher education, Research) | International | VALIDATED |  |
| 63 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Impact of altimetric data assimilation in a eddy-permitting, coupled physical-biochemical model of the North Atlantic ocean   | 28/10/2014 | Lake Constance, Germany (2014 OSTST meeting)                                  | Scientific community (higher education, Research) | International | VALIDATED |  |
| 64 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Toward variational assimilation of SARAL/AltiKa altimeter data in a North Atlantic circulation model at eddy-permitting resolution: assessment of a NEMO-based 4D-VAR system                | 28/10/2014 | Lake Constance, Germany (2014 OSTST meeting)                                  | Scientific community (higher education, Research) | International | VALIDATED |  |
| 65 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Reduced-order data assimilation methods   | 24/06/2013 | Bologna, Italy (Bologna Advanced Data Assimilation Summer School)             | Scientific community (higher education, Research) | International | VALIDATED |  |
| 66 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE      | Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modelling   | 16/03/2013 | University of Oxford, UK (Workshop)   | Scientific community (higher education, Research) | International | VALIDATED |  |

|    |   |   |   |            |   |               |           |  |
|----|---|---|---|------------|---|---------------|-----------|--|
| 67 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE                              | System evaluation   | 10/03/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 68 | Oral presentation to a scientific event | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | On the impact of nonlinearity on ensemble smoothing   | 08/04/2013 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 69 | Posters                                 | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | Localization in ensemble data assimilation  | 08/10/2012 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 70 | Oral presentation to a scientific event | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | Ensemble smoothing under the influence of nonlinearity  | 02/07/2013 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 71 | Oral presentation to a scientific event | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | Adaptive localization in ensemble Kalman filter methods by controlling the observation space  | 07/10/2013 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 72 | Oral presentation to a scientific event | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | Ensemble Data Assimilation: Algorithmic and Practical Aspects   | 21/11/2013 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 73 | Posters                                 | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | A comparison of linear and non-linear data assimilation techniques using a toy model  | 24/02/2014 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 74 | Oral presentation to a scientific event | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | The HBM-PDAF assimilation system for operational forecasts in the North and Baltic Seas   | 28/10/2014 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 75 | Posters                                 | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | Extending MEMO for Ensemble Data Assimilation on Supercomputers with the Parallel Data Assimilation Frame   | 29/01/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 76 | Posters                                 | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | Building Ensemble-Based Data Assimilation Systems   | 13/04/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 77 | Oral presentation to a scientific event | STIFTELSEN NANSEN SENTER FOR MILJOOG FJERNWALING                          | Extending the square root method to account for model noise in the ensemble Kalman filter   | 09/05/2013 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 78 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE                              | CNRS-LEGOS contribution to SANGOMA  | 26/11/2012 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 79 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE                              | Impact of observations (TD, travaux dirigés)  | 24/06/2013 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 80 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE                              | The SANGOMA Tools For Data Assimilation   | 22/09/2014 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 81 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE                              | Toward variational assimilation of SARAL/AltiKa altimeter data in a North Atlantic ocean model at eddy-permitting resolution: assessment of a NEMO-based 4D-VAR system                | 13/04/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 82 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE                              | Ensemble assimilation of ARGO temperature profile, sea surface temperature and Altimetric satellite data into an eddy permitting primitive equation model of the North Atlantic ocean | 13/04/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 83 | Oral presentation to a scientific event | CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE                              | Ensemble assimilation of JASON/ENVISAT and JASON/ARKA altimetric observations with stochastic parameterization of the model dynamical uncertainties                                   | 13/04/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 84 | Oral presentation to a scientific event | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | A comparison of linear and non-linear data assimilation methods using the NEMO ocean model  | 28/05/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 85 | Posters                                 | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | The SANGOMA Tools For Data Assimilation   | 04/05/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 86 | Posters                                 | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | A comparison of linear and non-linear data assimilation methods using the NEMO ocean model  | 13/04/2015 | Scientific community (higher education, Research) | International | VALIDATED |  |
| 87 | Posters                                 | ALFRED-WEGENER-INSTITUT HELMHOLTZ-ZENTRUM FUER POLAR- UND MEERESFORSCHUNG | Extending NEMO for ensemble data assimilation on supercomputers with the parallel data assimilation framework PDAF  | 27/04/2014 | Scientific community (higher education, Research) | International | VALIDATED |  |