REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*

2. REPORT TYPE
Conference Proceeding

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Assessment and Testing of the GODAE Products

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
C. Leprovost, M. Bell, E. Chassignet, James A. Cummings, I. Fukumori, Harley E. Hurlburt, Masafumi Kamachi

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Research Laboratory
Oceanography Division
Stennis Space Center, MS 39529-5004

8. PERFORMING ORGANIZATION REPORT NUMBER
NRL/PP/7304/02/0002

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217-5660

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

20040203 160

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The aim of this paper is to review how the GODAE centers will proceed to ensure the value and quality of their ocean products and to evaluate the performance of their systems. The strategy is to define a set of standard internal verification tests and metrics. The scientific credibility will rely on careful checks of the consistency of the system outputs with state-of-the-art knowledge of the ocean state and its variability. The quality assurance of the products will rely on systematic verification of key parameters and computation of statistical indexes by reference to both climatologies and real time data, and, in a delayed mode, to quality controlled observations. The performance of the systems will rely on diagnostics based on key indicators such as estimates of forecasting skill, ability to constrain a sparsely observed field or non-assimilated field, and evaluation of real time versus reanalysis products. A few examples of metrics for intercomparison will be given from operational systems that are specific to the Atlantic and the Pacific basins, where our knowledge of the ocean characteristics is the most advanced and where comparison exercises are under way as part of the GODAE common.

15. SUBJECT TERMS
GODAE, real time

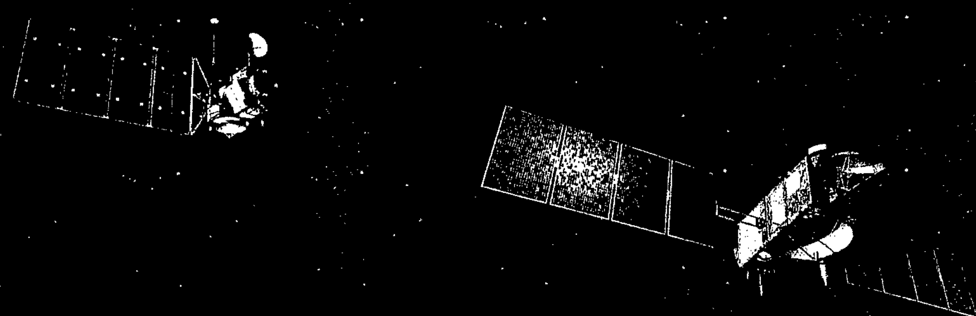
16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified	Unclassified	Unclassified	SAR	7	Harley Hurlburt
					19b. TELEPHONE NUMBER *(Include area code)* 228-688-4626

GODAE

Global Ocean Data Assimilation Experiment

cnes

CENTRE NATIONAL D'ETUDES SPATIALES



Symposium International / *International Symposium*

"En route to GODAE"

13 - 15 juin/*June* 2002, Biarritz, France

ACTES/*PROCEEDINGS*

Assessment and testing of the GODAE products

Christian LE PROVOST[1], Mike BELL[2], Eric CHASSIGNET[3], Jim CUMMINGS[4], Ichiro FUKUMORI[5], Harley HURLBURT[6], Masafumi KAMACHI[7]

[1] LEGOS / CNRS, Observatoire Midi Pyrénées, 14 Av. E. Belin, 31 400, TOULOUSE, France

[2]Met Office, London Road, Bracknell, Berkshire RG12 2SZ, United Kingdom

[3]RSMAS/MPO, University of Miami, Miami, FL, 33149, USA

[4]Naval Research Laboratory, Code 7533, Monterey, California 93943, USA

[5]Jet Propulsion Laboratory, Mail stop 300-323, 4800 Oak Grove Drive, Pasadena, California 91109, USA

[6]Naval Research Laboratory, Code 7304, Stennis Space Center, MS 39529-5004, USA

[7]Meteorological Research Institute, 1-1 Nagamine, Tsukuba 305-0052, Japan

Christian.Le-Provost@cnes.fr

Abstract – The aim of this paper is to review how the GODAE centers will proceed to ensure the value and quality of their ocean products and to evaluate the performance of their systems. The strategy is to define a set of standard internal verification tests and metrics. The scientific credibility will rely on careful checks of the consistency of the system outputs with state-of-the-art knowledge of the ocean state and its variability. The quality assurance of the products will rely on systematic verification of key parameters and computation of statistical indexes by reference to both climatologies and real time data, and, in a delayed mode, to quality controlled observations. The performance of the systems will rely on diagnostics based on key indicators such as estimates of forecasting skill, ability to constrain a sparsely observed field or non-assimilated field, and evaluation of real time versus reanalysis products. A few examples of metrics for intercomparison will be given from operational systems that are specific to the Atlantic and the Pacific basins, where our knowledge of the ocean characteristics is the most advanced and where comparison exercises are under way as part of the GODAE common.

1 – Introduction

The assimilation centers need to evaluate the performance and effectiveness of their systems. It is also useful to define relevant metrics measuring the quality of the products to aid users in assessing their usefulness. Work is underway to design quantitative evaluation methodologies for carrying out these assessments. This requires the definition of metrics which can be systematically used by the assimilation centers. Experience will be shared through critical analysis of intercomparison exercises. We will focus here on "internal metrics", i.e., the metrics considered by the assimilation centers to insure the value and quality of their products and to evaluate the performance of their systems. We defer discussion of "external metrics" that measure the impact of observing systems and assess assimilation products for different applications. More experience in product utilization is deemed necessary to define effective measures of external usefulness.

2 – Scope of Consideration

The quantities which will be considered in the internal metrics are primarily the model states that are functions of location M and time t. These include temperature T, salinity S, velocity in the u-east-west direction, and v-north-south direction, time-varying mixing tensors, sea surface height, and other passive tracers. From these variables, other quantities are derived that are representative of major dynamical and thermodynamical ocean characteristics, such as vertical velocity, volume transports of major currents, mixed layer depth and vertical temperature/salinity profiles, thermocline depth, heat fluxes, potential vorticity, and water mass characteristics. The significance of the state under consideration will depend on processes resolved by the analysis and forecasting systems. In the discussion below, it is important to recognize what is resolved by the products and what is not. Differences between model estimates and reality are due to inaccuracies in what the models resolve on the one hand and incompleteness of the model in its ability to represent aspects of the real world on the other. The latter element does not indicate significance of the former, or vice versa, and metrics must be interpreted accordingly. In particular, accurate descriptions of certain aspects of ocean circulation (resolved space) can be valuable even though it may fail to describe other aspects of the ocean (unrepresented space). Such distinctions may be found in differences in resolved space- and time-scales such as related to the level of priority in respective assimilation systems: sea surface and mixed layer versus upper thermocline or full depth, mesoscale versus large scale (the ocean weather scale versus the ocean climate scale), high frequency -including inertial gravity waves- versus low frequency – monthly to seasonal-to-interannual –, upper ocean dynamics and thermodynamics versus deep ocean circulation and climate. Distinctions may also be found in the geographical extent of the systems: regional scale (western boundary currents, straits, frontal zones, subduction or convection areas), basin scales, and global scale.

3 – GODAE strategy for internal assessments

The internal metrics can rely on several types of diagnostics. One approach is to consider the following classification: 1) assessments based on the experience in understanding and modeling of the ocean by the oceanographic community: we call this class of metrics "consistency analysis", 2) assessments based on direct comparisons with available observations of some of the above listed variables, accessible either in real time, or delayed mode: we call this "quality analysis", and 3) assessments aiming at the evaluation of the technical effectiveness of the systems in terms of modeling and data assimilation strategies: we call this "performance analysis". It is important to notice before going further that these diagnostics can be applied on quantities issued either from free mode model runs, or from analysis/forecasting systems including assimilation. Diagnostics on free mode runs are of major interest for evaluating the strengths and weaknesses of the models. Diagnostics on the assimilated products will also permit evaluating the added value of the data assimilation process.

3.1 Consistency analysis: diagnostics based on our scientific understanding of the ocean

The scientific credibility will rely on a careful check of the consistency of the system outputs with state-of-the-art understanding of the ocean state and its variability. In particular, verifications will focus on the major well known possible weaknesses of the modeling and assimilation systems, and the systematic errors often arising from the variety of models, physical parameterizations and assimilation schemes. Ocean modeling has now reached a state of considerable maturity. The realism of model outputs is continuously improving, through better parameterization of the ocean physics, better forcing inputs, better resolution, better numerical schemes, and increased computational capabilities. Assimilation of observations in these models results in still better estimates of the ocean state and its evolution. The evaluation of this realism is necessarily limited because ocean observations are generally insufficient in space and time. It generally relies on our up-to-date knowledge of the ocean state based on synthesis of historical data sets and theoretical understanding. For the following, we must acknowledge reference in particular to the DAMEE-NAB (Chassignet and Malanotte-Rizzoli, 2000) and the DYNAMO (Meincke, Le Provost and Willebrand, 2001) special issues which reported on two major inter-comparison experiments of free run simulations over the North Atlantic Ocean.

3.1.1 Time-Mean Climatologies

The zero order control of consistency must address the mean statistically steady state of the simulated ocean fields, on the horizontal and vertical dimensions. Of course, there is a priori a difficulty on the definition of the "mean" state. The ocean spectrum is red, and, strictly speaking, the concept of a "steady ocean state" is not well founded. To be pragmatic, we suggest to adopt as a rule to compute "time-means" over the longest available time series of simulated products, with at least one year duration, in order to eliminate the seasonal signal.

3.1.1.1 Sea surface fields: we have some robust knowledge of the characteristics of these fields from the compilation of observations (from in situ and from remote sensing). This includes temperature, salinity, and sea level. The actual characteristics of these mean fields are generally excessively smooth compared to what must be the reality, because of the limitations in the way we can observe them. But these means are indicative of the global behaviour of the simulations, in terms of spatial position of the major fronts and ocean currents.

3.1.1.2 Integrated transports through sections: volume transports through particular sections are major indicators of the realism of the outputs. This includes checking of the order of magnitude of the transports through straits, like the Florida Strait, Gibraltar, Drake, or in the major Western Boundary Currents.

3.1.1.3 Vertical structure of the major current systems: the three dimensional structure of the ocean circulation results in complex vertical distribution of the ocean flows, whose major typical features are more or less well known from often repeated hydrographic sections. One typical example is the vertical structure of the Atlantic thermohaline circulation.

3.1.1.4 Water mass characteristics in temperature and salinity: the analysis of the ocean mass properties is a key diagnostic of the system's behaviour. Models are initialized with climatologies. In regions in contact with the atmosphere, the water mass properties are determined by the atmospheric forcing and hence deviate from the initial conditions. Incorrect characteristics in water masses may reveal problems in the forcing fields, or surface boundary layer parameterization. In the deeper regions that are not ventilated on short to medium term, deviations from the climatologies may indicate model problems.

3.1.1.5 Transports in density classes: the analysis of volume transports in density or temperature classes is a powerful diagnostic of the dynamics and the mixing in model simulations. One typical example is the analysis of the simulated overflows in the three DYNAMO test models across the Iceland-Scotland Ridge and the Denmark Strait.

3.1.1.6 Sea ice distribution: accurate knowledge of sea ice distribution is important for several reasons. First, sea ice serves as an effective insulation between the cold atmosphere and relatively warmer ocean. Errors in the sea ice distribution used as a lower boundary condition in the atmospheric model can affect flux estimates of heat out of the ocean by orders of magnitude. Second, sea ice affects the surface albedo and can effectively eliminate solar radiation as a source of heating into the ocean in ice covered areas. Finally, sea ice distribution affects ocean circulation directly by increasing the density of the ocean water under the ice, inducing convective processes that deepen mixed layers and ultimately contribute to driving the deep thermohaline circulation of the ocean in regions of overturning. On smaller time and space scales, the location of the ice edge influences atmospheric forcing through cyclogenesis of intense polar lows. Uncertainty in the distribution of sea ice on these scales results in errors in the winds used to force the ocean model.

3.1.1.7 Thermohaline Circulation: the classical diagnostics include the estimate of the Meridional Overturning Circulation and of the Meridional Heat Transport. A correct MOC and heat transport are indicative of the model's performance. It also offers a first assessment of the water mass transformations that take place within the model. For the Atlantic, the MOC strength can directly be related to western boundary current strength

3.1.2 Space and Time Variability

The ocean is highly turbulent and also shows a large range of variability on all space- and time-scales due to its internal physics and to the variability of the external forcings (mainly from the atmosphere). Our knowledge of the characteristics of this variability is far from being complete. However, the impressive amount of work completed since the 70's gives us some bounds on what the relevant space and time scales of the ocean variability are. And the progress in remote sensing of the ocean, and use of autonomous in situ sensors (surface drifters, floats) is progressively improving our knowledge in the field. By reference to the literature, the following basics diagnostics must be considered.

3.1.2.1 The upper ocean mixed layer: the mixed layer characteristics are of major interest for many applications ranging from weather and seasonal predictions or biological investigations to applications for military needs and fisheries. They are highly variable, in space and time. The usual diagnostics are on the space scales of the mixed layer depth and its variability, SST, and on the estimates of the ocean-atmosphere heat and momentum fluxes.

3.1.2.2 The variability of the surface ocean currents: some current systems are known to behave within some boundary limits which are possible to check in the model outputs. This is the case for the Gulf Stream meandering, for the Kuroshio bi-modal behaviour, and for some current systems highly subject to seasonal variations, such as the North Brazil current.

3.1.2.2 The statistics on the eddy field: comparisons on the variability statistics (geographical distribution, amplitude, frequency spectra, spatial scales) of the model fields can be made with similar quantities obtained on long time span from altimetry, ocean colour, surface drifters. The mapping of the sea surface height variability, and of the near surface eddy kinetic energy is now a standard diagnostic that can be applied globally for any eddy resolving ocean model and analysis and prediction system. More specific diagnostics are related to limited locations where vertical or abyssal eddy kinetic energy distribution has been obtained, on short time span, from in situ current meter moorings.

3.1.3 Physical Balance

A third element of consistency concerns relationships among model state variables with respect to our theoretical understanding of ocean circulation. These include various balances among variables at a given instant (including the time-mean) and the temporal evolution of the model state. While specific details depend on the nature of the models employed (e.g., quasi-geostrophy versus primitive equation), a description of the ocean that is physically consistent is a necessary element in understanding ocean circulation and its changes.

3.1.3.1 *Instantaneous balance.* To first approximation, large-scale velocity fields of the ocean are in geostrophic balance away from the Equator. Along the Equator, a second order balance generally exists between zonal velocity and the meridional pressure gradient. Velocity normal to topography and the coast should trivially vanish and the three-dimensional velocity field should be non-divergent. Temperature and salinity fields should be such that the water column is statically stable. Apart from low-frequency changes in the ocean's heat and salt content, the time-mean divergence of heat and salt (fresh water) fluxes, including advective and diffusive components, should be zero everywhere except at the surface where it should equal their respective external forcing. Similar balances should hold for other passive tracer fluxes.

3.1.3.2 *Temporal evolution.* Temporal differences among model state variables should be consistent with the state's implied advective and diffusive effects and external forcing. To the extent that there are no internal sources or sinks, temporal changes in heat and tracer content (including salt) should equal convergence of their respective fluxes at all space- and time-scales. Changes in circulation should be dynamically compatible with changes in available potential energy and external forcing. Potential vorticity should be conserved following a water parcel away from direct forcing and dissipative regions.

3.2 Quality assurance: diagnostics based on observations

The quality control of the products must rely on observations. These diagnostics will be necessarily limited by the availability of the observation data sets. They include XBT and SSS lines, time series of hydrographic sections, moorings, ADCP, sea level gauges, satellite SSH altimetry, satellite SST, drifters, profiling floats. Some of the data will be available in real time, others will not, resulting in a two level evaluation procedure: a real time loop based on systematic verification of key parameters and computation of statistical indexes by reference to real time data, and a delayed mode loop involving comparisons with quality controlled observations. Direct comparison of simulations to observations at mesoscales is essential for eddy-resolving data-assimilative models and their forecasts, although such comparisons will be of limited utility for ocean simulations run entirely in free mode, because of the chaotic aspects of mesoscale eddies. For the seasonal part of the variability forced by the atmosphere, such comparisons are feasible, provided the turbulent part of the signal is ignored.
For simplicity, we will consider in the following only the qualitative evaluation of products issued from systems including data assimilation. It must also be noticed that some of these observations will be assimilated in the systems. Their use appears thus at many steps: a priori in the assimilation step, and **a posteriori** in a diagnostic way at the validation step – this is what is presented in the following sections -, but also in a control of performance way, as will be presented in section 3.3.

3.2.1 Time Series Stations

In situ time series stations supply data which are directly usable for checking the accuracy of the system analyses (in real time or delayed mode), and the forecasts (a posteriori). It must be noticed that they imply "point comparisons" which have to be accommodated with the discrete sampling of the model outputs. Deviations from the observations can easily be computed, including evaluations of systematic biases. The main time series presently available include the following.

3.2.1.1 Barotropic transport: there exists at least one example of transport monitoring from phone cable measurement: the Florida Strait volume transport (Larsen, 1992). Daily transports are available from March 1982 to October 1998, and from March 2000 onward. The data are available on the PMEL and AOML NOAA ftp sites.

3.2.1.2 Sea level gauges: Sea level gauge measurements are available in many locations, easily accessible thanks to GLOSS at its fast delivery Center of Hawaii, or on its delayed mode archiving Center of Bidston (GLOSS implementation Plan, 2000). Up to recently, "fast" was meaning with a delay of typically a month. Now, it is improving, and for about one hundred stations, the delay will be no more than a week or less. The interest of sea level data, by comparison to altimetry, is the high frequency sampling of the order of one hour or less.

3.2.1.3 Moorings: for the time being, the main in situ networks of permanent moorings are located in the Equatorial Pacific (TAO-TRITTON) and the Equatorial Atlantic (PIRATA). These networks supply real time observations on atmospheric parameters in the atmospheric boundary layer, which can be used to test locally the surface fluxes included in the systems to force the ocean model, and on temperature, salinity, and horizontal velocity in the upper layer of the ocean, which allow to implement local diagnostics on the products delivered by the analysis/forecasting systems.

3.2.2 Satellite Remote Sensing

Remote sensing data coming from satellites are a major sources of information which have the great interest of being almost synoptic in space and time. They include mainly SSH altimetry and SST, but also other important data sets such as surface radiation, sea ice products, ocean color, and ocean bottom pressure. Remotely sensed salinity could also become available during the GODAE timeframe. These observations will be used in different ways. We have already pointed out in section 3.1.2.3 the use of satellite altimetry for statistics on the space and time variability of the sea surface topography and surface geostrophic currents. But these data sets can also be used to check along time the quality of the related products. Deviations in space and time (biais, rms difference, variances) can be considered which are useful indicators of the behaviour of the systems along time. We will come back on the use of these fields in the following in section 3.3. We must also notice here that these remote sensed data sets are however, up to now, not at the right sampling and accuracy for some global climate or for high resolution applications. IR and ocean color imagery can be useful in assessing the ability of eddy-resolving ocean prediction systems to map and forecast the position of individual mesoscale features.

3.2.3 VOS and Floats

Measurements from VOS and floats provide valuable observations inaccessible to other means above. The nature of these measurements is no different from other in situ observations except for their irregular sampling characteristics.

Comparisons of these measurement types with models can easily be made by extracting model output according to the space and time sampling of the VOS and float measurements. However, the irregular sampling characteristics of these measurements are not immediately amenable to separating anomalies from the time-mean, making characterization of prior observation errors somewhat difficult. (Climatological means and areal averages are often substituted as the time-mean reference.) There is some disagreement on the best usage of measurements of float displacement of whether to assimilate them as Lagrangian trajectories or converting displacements to average Eulerian velocities. The unique information content of Lagrangian trajectories with respect to Eulerian velocities would seem limited due to the trajectories being chaotic in the sense that they are extremely sensitive to unresolved small-scale flow; i.e., Lagrangian trajectory information may be dominated by model representation errors that cannot be utilized.

3.3 Performance: diagnostics based on statistical measures

In contrast to the diagnostics above, the third class of metrics, which we call "performance analysis", aims to evaluate the technical effectiveness of the assimilation systems. The measure involves the use of key indicators such as estimates of formal errors, forecasting skill, ability to constrain a sparsely observed field or estimate a non-assimilated field. In turn, performance metrics are also relevant to measuring consistency and quality discussed above. Examples of these statistical metrics can be found in Fukumori et al. (1999).

3.3.1 Error Estimates

Formal error estimates provide quantitative measures of accuracy and significance. Of primary interest are error covariances among the model state variables and those of the data constraints. The former measures accuracy and dependency of what is resolved whereas the latter includes model representation errors that are also model dependent as well as instrumental errors of the observing systems. With suitable linearization, errors of any model diagnostic variable can be derived from the model state error covariance matrix. The error estimates also provide measures of statistical consistency of the assimilation systems. Differences between model products and observations should be comparable with respective formal uncertainty estimates that are based on first principles. Differences in model product errors as functions of assimilated observations provide measures of the different observations' impact on estimation and can be used to design effective observing systems (observing system simulation experiments, OSSEs.)

3.3.2 Model-Data Differences

The accuracy of data assimilated model products is theoretically a non-decreasing function of the amount of data that is assimilated. A degradation caused by assimilation generally indicates inaccurate assumptions in the assimilation scheme. While models can be forced to agree with observations (e.g., replacing equivalent model fields with data), improvements with respect to independent observations are not trivial. An assessment of model improvement (or lack of degradation) with respect to non-assimilated, independent measurements is therefore an effective means of assessing the performance of an assimilation system. Variances of model-data differences serve as common measures of the estimates' accuracies. In particular, the simulation (non-assimilated, free model run) equivalent of the metrics below serves as the relative measure of this improvement and the assimilation's success.

3.3.2.1 *Innovation Vector.* The innovation vector is the difference between observations that are about to be assimilated and the prior guess by the model. The innovation vector is routinely evaluated in sequential assimilation schemes (difference between data and model forecast) and its variance provides a readily available measure to monitor the effectiveness of the assimilation system.

3.3.2.2 *Residual Vector.* The residual vector is the difference between observations and the prior guess after the observations have been assimilated. In principle, the analysis residuals should be spatially uncorrelated. Any spatial correlation remaining in the residuals represents information that has not been extracted by the assimilation system. By stratifying the residuals by observing system, information on how effective an observing system is being utilized and observing system biases can be determined.

3.3.2.3 *Forecasting skill.* Observations that formally lie in the future provide an independent set of data to assess the assimilation system. Forecasting skill (i.e., differences between observations and a simulation initiated from an assimilated state using data prior to the compared observations) is a common metric used in numerical weather prediction and can be employed as an effective posterior measure for any assimilation and forecasting system. The innovation vector above is in essence a forecasting skill, but one that is limited to the data sampling period

3.3.2.4 *Withheld Observations.* Withholding parts of the observations and using them as independent measures of accuracy allows a direct testing of the goodness of the assimilation system. In particular, withholding certain classes of observations that are most independent of those that are assimilated is effective in assessing and optimizing assimilation of different observation types; e.g., withholding subsurface measurements in assimilating satellite remote sensing of the sea surface. However, once tested, optimality requires all available observations to be assimilated, and continually withholding independent observations is not desirable.

3.4 Visual Evaluation

Visual means can be useful in evaluating the ability of an ocean prediction system to represent and forecast ocean features of interest as well as in detecting temporal oscillations, unphysical results and numerical noise. This includes animations which, for example, can show oscillations, trackiness in assimilation of satellite altimeter data and mesoscale eddies that unphysically wax and wane in sequence of analyses.

4 Pilot intercomparison experiments

Judging the strength and weakness of each system, identifying errors and their origins, and clarifying the value of sophisticated assimilation schemes and parameterization of physical processes, are difficult and laborious tasks. The sharing of experience is thus critical. Inter-comparison exercises between the different GODAE Centers is one way to respond to this need. Such exercises however are not easy, as was illustrated for ocean modeling by the recent DAMEE-NAB and DYNAMO experiences. Two strategies are possible: controlled experiments, in which key elements of the systems are fixed, such as domain, grid resolution, boundary conditions, forcing fields, ... and free mode experiments in which the constraints agreed to are only the area and the period of the exercise. The inter-comparison goals can also involve several levels: delayed (research/reanalysis) mode for assessing the performance of the integrated systems, or real time (operational) mode for assessing the feasibility of the real time analysis/forecasting systems. The metrics for conducting these inter-comparisons are a subset of the extended list presented in section 3. Two GODAE Pilot Projects for inter-comparisons have been started.

4.1 The North Atlantic case

The use of the Atlantic as a prototype domain to test and evaluate how practically an inter-comparison exercise can be carried within GODAE has been agreed because of the state of development of the different components of an ocean forecasting system over this basin: already well instrumented, large number of available models, high user interest. A pilot project has been initiated through the INTERCAST proposal (DeMey et al, 2001) agreed between the FOAM and MERCATOR forecasting systems. Other groups including HYCOM, NCOM, and NLOM have expressed their willingness to join the exercise (see GODAE implementation plan). The main characteristics of this inter-comparison exercise are the following.

1. The exercise will focus on the North Atlantic and will cover the period January 2000 - July 2001
2. The exercise will consist in comparing similar diagnostics and fields for similar simulations of each system
3. Integrations of the models will be performed and assessed chiefly with assimilation of observational data
4. At least one integration will be performed by each group in which a similar subset of observations (namely, satellite altimetry) are assimilated
5. The surface fluxes, assimilation data and general procedures used to drive the systems will be those used by the systems for real-time analyses
6. A core set of diagnostics will be agreed together by both project teams, following an initial recommendation (Le Provost, 2001).
7. The intercomparison exercise will cover (i) analyses (hindcasts); (ii) 7- and 14-day-range forecasts with analysed or forecast atmospheric forcings. The form of the diagnostics will be time series (of spatial averages when needed), and fields (of temporal averages when needed). At least an annual cycle will be covered by all diagnostics, except forecast diagnostics which will be calculated in specific periods of the year (at least: February 15 - March 15 and August 15 - September 15).

This project is scheduled to finish by the end of 2002. It will be complemented by a new initiative, MERSEA (Marine EnviRonment and Security for the European Area), funded by the Eureopean Community, in a wider context including assessments not only on the operational model systems, but also on the operational observation network, and demonstration of different system application from user perspectives.

4.2 The North Pacific case

Intercomparison exercise will give experience and information on how the GODAE centers will proceed to evaluate and ensure the quality of assimilation products and systems. It will, as a result, promote the international GODAE. The Japan-GODAE working team, therefore, proposed a North Pacific intercomparison project. The intercomparison project has been agreed to in the IGST meetings. A set of metrics in the North Pacific was reported and discussed at the "International Workshop on GODAE with Focus on the Pacific", IPRC, July 2001 (Kamachi and Minato, 2001; see also the GODAE Implementation Plan). A similar pilot project of comparison between assimilation products and observations has been initiated in the Japan Meteorological Agency (JMA) in 2001. With these experiences, the Japan-

GODAE working team has initiated a North-Pacific intercomparison project in 2002 in cooperation with GODAE partners in the USA and Asian countries. The main characteristics of this intercomparison exercise are the following.

1. The exercise will focus on the North Pacific and will cover the period January 2000 - December 2001
2. The exercise will consist in comparing similar diagnostics and observation fields. The information will be delivered from IPRC and JMA.
3. Integrations of the models will be performed and assessed chiefly with assimilation of observational data (same as the North Atlantic case).
4. At least one integration will be performed by each group in which a similar subset of observations (namely, satellite altimetry) are assimilated (same as the North Atlantic case).
5. The surface forcings (fluxes), assimilation data and general procedures used to drive the systems will be those used by the systems for real-time operational (or delayed mode research) analyses. It is a kind of free experiment and similar to the North Atlantic case.
6. A set of diagnostics will be agreed by Japan GODAE working team and collaborators, following an initial report by Kamachi and Minato (2001).
7. The inter-comparison exercise will cover (i) reanalyses (hindcasts); (ii) 7-, 14- or 30-day-range forecasts with analysis or anomaly added climatology of atmospheric forcings. The form of the diagnostics will be 2D fields and time series. Annual, monthly or shorter variability (or of specific period) will be covered (similar to the North Atlantic; see also GODAE Implementation Plan).
8. The products of each partner will be submitted to the IPRC data center.

This project is scheduled to be finished by the end of 2002 (same as North Atlantic case). A report will be submitted to the IGST Meeting.

5 Conclusions

The synergy between the different GODAE Centers is critical for insuring the success of the Experiment and in particular for improving the effectiveness and quality of the different systems. The sharing of a common strategy for assessing the performance of the systems and testing the quality of the outputs is an important component of the GODAE Common. A consensus on the definition of a standard set of internal metrics and on their systematic use is a first step. A rationale has been proposed here to build this list, which is not exhaustive and which calls for enrichment based on on-going developments. The inter-comparison exercises planned for the North Atlantic and the North Pacific will be one way to build on this needed close relationship. Preliminary results are presented in the poster session of this Symposium.

References

[1] E. Chassignet and P. Malanotte-Rizzoli, DAMEE-NAB , *Dyn. Atmosphere Oceans*, Vol 32, N°3-4, 2000.

[2] J. Meincke, C. Le Provost and J. Willebrand. Dynamics of the North Atlantic Circulation (DYNAMO). *Progress in Oceanography*, Vol. 48, N°2-3, 2001.

[3] J.C. Larsen, Transport and heat flux of the Florida Current at 27°N derived from cross-stream voltages and profiling data : theory and observations, *Philosophical Transactions of the Royal Society of London*, A 338, 169-236, 1992.

[4] GLOSS implementation Plan, 2000.

[5] P. DeMey, M. Bell, E. Greiner, M. McCulloch, C. Le Provost, Intercomparison of Outputs from Atlantic Ocean Forecasting Systems, INTERCAST, *FOAM/MERCATOR Internal report*, December 2001.

[6] C. Le Provost, GODAE Metrics over the Atlantic Ocean. *GODAE internal report*, December 2001

[7] I. Fukumori, R. Raghunath, L. Fu, and Y. Chao, 1999. Assimilation of TOPEX/POSEIDON data into a global ocean circulation model: How good are the results?, *Journal of Geophysical Research*, 104, 25,647-25,665.

[8] M. Kamachi and S. Minato, Pacific Inter-Comparison (Internal Metrics), *Handout in the International Workshop on GODAE with Focus on the Pacific*, IPRC, Hawaii, July 2001