

## Observing System Evaluation

Peter R. Oke<sup>1</sup>, Magdalena A. Balmaseda<sup>2</sup>, Mounir Benkiran<sup>3</sup>, James A. Cummings<sup>4</sup>, Eric Dombrowsky<sup>3</sup>, Yosuke Fujii<sup>5</sup>, Stephanie Guinehut<sup>6</sup>, Gilles Larnicol<sup>6</sup>, Pierre-Yves Le Traon<sup>7</sup>, Matthew J. Martin<sup>8</sup>

<sup>1</sup>CAWCR/CSIRO and WfO National Research Flagship, Hobart, Tasmania, Australia

<sup>2</sup>European Centre for Medium-Range Weather Forecasting, Reading, United Kingdom

<sup>3</sup>Mercator Ocean, Ramonville Saint Agne, France

<sup>4</sup>Oceanography Division, Naval Research Laboratory, Monterey, USA

<sup>5</sup>Oceanographic Research Department, Meteorological Research Institute, Tsukuba, Japan

<sup>6</sup>CLS-Space Oceanography Division, Toulouse, France

<sup>7</sup>Ifremer, Technopole de Brest-Iroise, Plouzane, France

<sup>8</sup>Met Office, Exeter, United Kingdom

Corresponding author: Peter Oke, CSIRO Marine and Atmospheric Research, GPO Box 1538, Hobart 7001, Tasmania, Australia ([Peter.Oke@csiro.au](mailto:Peter.Oke@csiro.au))

### Abstract

Global ocean forecast and reanalysis systems, developed under the Global Ocean Data Assimilation Experiment (GODAE), are a powerful means to assess the impact of different components of the Global Ocean Observing System (GOOS). GODAE systems can be exploited to help identify observational gaps and to ultimately improve the efficiency and effectiveness of the GOOS for constraining ocean models for ocean prediction and reanalysis. Many tools are currently being used by the GODAE community to evaluate the GOOS. Observing System Experiments (OSEs), where different components of the GOOS are systematically with-held, can help quantify the extent to which the skill of a model depends on each observation type. Observing System Simulation Experiments (OSSEs), where the impact of hypothetical observation types is assessed through some form of twin experiment, help evaluate the potential benefits of future observation types. Singular vector analysis can quantify the sensitivity of the ocean circulation to small perturbations, potentially identifying regions where additional observations are likely to benefit ocean forecasts. Analysis sensitivity, via the influence matrix, and adjoint-based techniques can be used to routinely assess the influence of individual observations and observation types on an analysis. These methods are therefore useful for evaluating and improving assimilation systems and for identifying the most and least beneficial components of the GOOS for a given system. A suite of examples using these approaches to observing system evaluation from a GODAE perspective are presented in this paper.

**Key words:** Data Assimilation, Observing Systems, Ocean Prediction

### 1. Introduction

The development and application of data assimilation techniques for numerical weather prediction (NWP) and more recently for operational oceanography have led to the growing use of models and assimilation tools for the assessment and design of atmospheric and oceanic observing systems (e.g., Rabier et al. 2008). The development of operational ocean forecast systems is a key initiative of the Global Ocean Data Assimilation Experiment (GODAE). All GODAE systems are underpinned by the Global Ocean Observing System (GOOS; [www.ioc-goos.org](http://www.ioc-goos.org)) that is comprised of satellite altimetry, satellite sea surface temperature (SST) programs, delivered through the GODAE High Resolution SST effort (GHRSSST; [www.ghrsst-pp.org](http://www.ghrsst-pp.org)), and in situ measurements from the Argo program, the tropical moored buoy, XBT and tide gauge networks. Each of these observation programs are expensive and require a significant international effort to implement, maintain, process and disseminate. While many components

of the GOOS are primarily intended for climate applications, their application to operational ocean forecast systems, such as those developed under the GODAE, is important. In this paper, we present results from Observing System Experiments (OSEs), Observing System Simulation Experiments (OSSEs), and diagnostics from assimilation systems that seek to assess the benefits of different observation types and arrays to realistic ocean forecast and reanalysis systems.

The OSEs described here generally involve the systematic denial, or with-holding, of different observation types from a data assimilating model in order to assess the degradation in quality of a model forecast or analysis when that observation type is not used. These OSEs seek to quantify the impact of different observation types on each specific forecast or analysis system. Importantly, the impact of each observation type may strongly depend on the details of the model into which they are assimilated, the method of assimilation and the errors assumed at the assimilation step. We therefore present results from a range of different models and applications in an attempt to identify the robust results that are common to a number of different systems.

OSSEs typically involve some sort of twin experiment, where “synthetic observations”, usually extracted from a model, are assimilated into an alternative model or mapped using an analysis system. Similarly, variational methods, such as singular vector and normal mode analysis, do not use real observations, but instead interrogate model physics and model sensitivities to identify regions in which small perturbations are quickly amplified. It is assumed that additional observations in regions of high sensitivity are likely to better constrain a data assimilating model. These types of analyses, though idealised, may be used to assess the impact of some hypothetical array of observations that may not exist yet. This means that OSSEs and variational methods can be used to contribute to the design of future observing systems, quantifying their possible impacts and limitations.

Techniques that are increasingly being used by the NWP community, for assessing and monitoring observing systems, include analysis sensitivity (e.g., Cardinali et al. 2004) and adjoint-based procedures (Langland and Baker, 2004). These techniques effectively diagnose the impact of assimilated observations on a given analysis or forecast. Though not widely adopted by the ocean forecasting community yet, several groups are developing capabilities in this area, with the intention of routinely applying them to operational systems to monitor and assess the time-evolving GOOS.

The inaugural Ocean Observing Panel for Climate (OOPC) - GODAE meeting on OSSEs and OSEs was held at UNESCO/IOC in Paris, France in November 2007 ([www.godae.org/OSSE-OSE.html](http://www.godae.org/OSSE-OSE.html)). This was the first international meeting dedicated to the subject of observing system evaluation using GODAE systems. Many of the ideas and results presented in this paper are based on presentations from the OOPC-GODAE OSSE/OSE meeting. This paper is organised as follows: results from various OSEs and OSSEs are presented in sections 2 and 3 respectively; a description of emerging techniques is included in section 4 and some conclusions are reported in section 5.

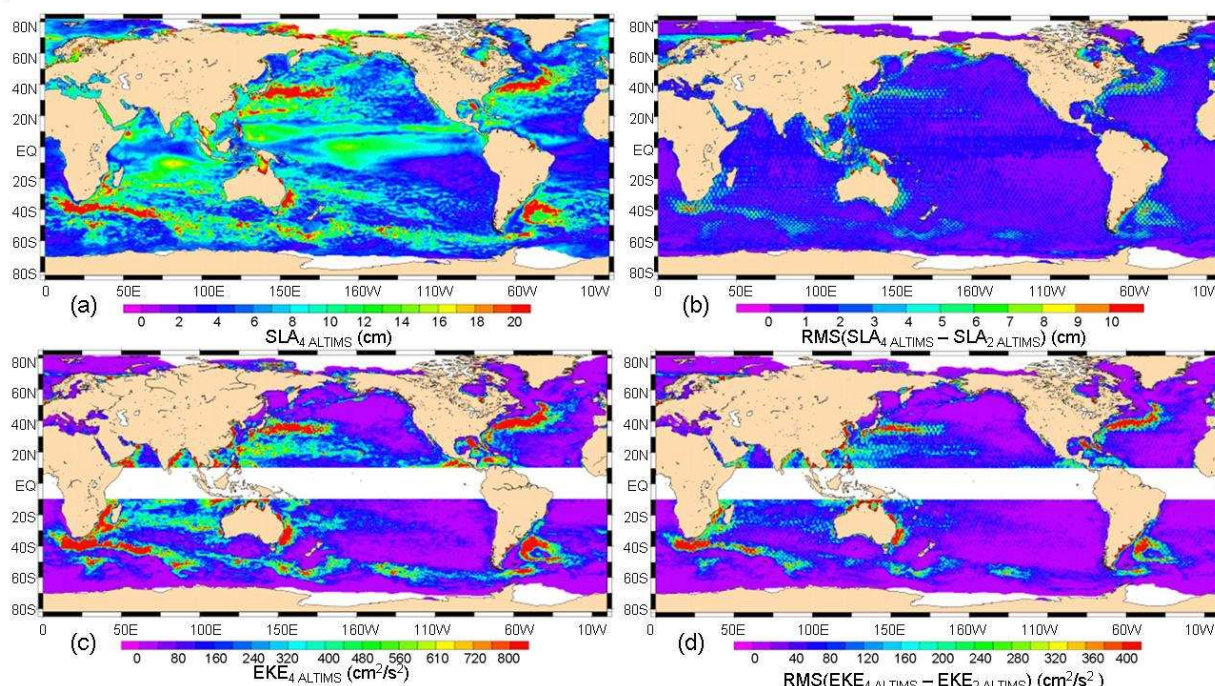
## **2. Observing System Experiments (OSEs)**

A determination of the requirements of the GOOS for operational oceanography is the primary goal of the studies described in this section. Collectively, we seek to assess the importance of different observation types for meeting the needs of a variety of operational oceanographic applications, ranging from observation-based mapping systems, like that of CLS/Aviso, short-range prediction systems from GODAE partners (e.g., Bluelink, Mercator, NRL, UK Met Office), and seasonal prediction (e.g., ECMWF). These studies also seek to assess the coverage of observing systems that is needed for different applications: the number of satellite altimeters is a recurring question that several groups have considered.

### *a. Observation-based Mapping Systems*

The Ssalto/Duacs center has led several studies aiming to identify the most appropriate satellite configuration to observe the mesoscale ocean variability using an analysis system, rather than a modeling system. Focusing first on the Mediterranean Sea, and later on the global oceans, Pascual et al. (2006;

2007) have demonstrated the benefits of merging data from four altimeter missions to produce high resolution global and regional maps of sea level anomalies (SLA). An illustration of the impact of the number of altimeters on mapped SLA and eddy kinetic energy (EKE) using the Ssalto/Duacs processing system (based on Optimal Interpolation, OI; Ducet et al. 2000) is given in Figure 1. Here the EKE is computed from geostrophic velocities that are derived from gridded SLA. In areas of intense variability, the root-mean-squared (RMS) differences between a classical configuration of two altimeters (Jason-1+ERS2/Envisat) and the scenario merging data from four altimeters can reach 10 cm for SLA and 400  $\text{cm}^2/\text{s}^2$  for EKE. This represents a significant percentage of the signal variance. At mid- and high-latitudes, previous studies have shown a clear underestimation of EKE due to the high frequency and high wavenumber signals that cannot be adequately mapped by combining data from only two altimeters (Ducet et al. 2000; Le Traon and Dibarboure 2002; Brachet et al. 2004).



**Figure 1: (a, c) RMS of SLA and near-surface EKE estimated with data from 4 altimeters (Jason-1 + T/P interleaved + ERS2/ENVISAT + GFO); (b, d) RMS difference between mapped SLA (and near-surface EKE) using data from 4 and 2 altimeters; adapted from Pascual et al. (2006).**

The impact of four altimeters is expected to be particularly important for operational forecast and analysis systems. Pascual et al. (2008) quantify the degradation in the quality of the altimeter products when Near-Real-Time (NRT) data are used compared to when Delayed-Time (DT) data are used. Three main sources of errors are identified in NRT data: the orbit is less accurate; the latency of data is a problem; and observation windows necessarily favour “old” data for NRT systems. Validation with independent in-situ data demonstrates the degradation of NRT maps compared to DT maps (Table 1 and

Table 2). That is, 4 altimeters in NRT are needed to get the same performance as 2 altimeters in DT. The statistics in Table 1 show comparisons between near-surface velocities derived from drifting buoys and SLA maps. Similarly, the statistics in

Table 2 show comparisons between SLA from tide gauges and SLA maps. Some of the discrepancies between SLA tide gauge data and SLA maps are probably due to the mapping procedure in addition to the errors in the altimeter data and to the number of altimeters used. For both tables, SLA maps using 2 and 4 satellites, and using NRT and DT data are shown. The results from

Table 2 also highlight the importance of continuous advances in the processing of altimeter data. For instance the differences between the old and new DT data relates to specific corrections such as tidal and high frequency corrections. These improvements have resulted in a significant reduction in the errors of altimeter observations.

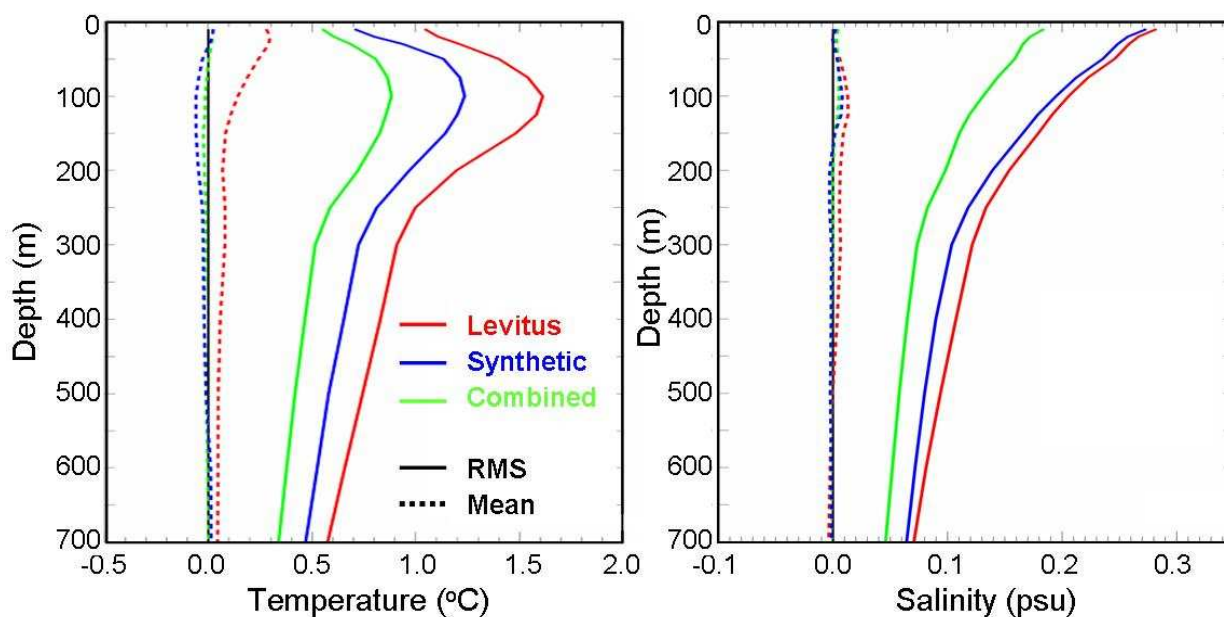
**Table 1: RMS difference (in cm/s) between drifter and altimeter-derived velocities in areas of intense variability (with RMS variability > 20 cm/s) and at latitudes of greater than 10° from the equator; adapted from Pascual et al. (2008). Comparisons are for the period October 2002 – August 2003.**

Variable	Delayed-Time		Near-Real-Time	
	2 missions	4 missions	2 missions	4 missions
U	11.23	10.72	12.06	11.30
V	10.70	9.97	11.63	10.69

**Table 2: RMS difference (in cm) between tide gauge sea-level and mapped altimetry for the old delayed-time, the new delayed-time and the near-real-time system; adapted from Pascual et al. (2008). Comparisons are for the period October 2002 - August 2003.**

Variable	Old Delayed-Time	New Delayed-Time	Near-Real-Time
2 missions	4.72	4.26	4.82
4 missions	4.27	3.94	4.42

A series of OSEs using the Global Observed Ocean Products (Larnicol et al. 2006) that combine remotely-sensed (SLA, SST) and in situ observations, using the method described by Guinehut et al. (2004; see section 3, below), facilitates a quantitative assessment of the relative contributions from different components of the GOOS. Figure 2 shows the RMS errors of sub-surface temperature (T) and salinity (S) using this approach (errors are based on differences between the mapped fields and in situ profiles that are used to produce the maps). This demonstrates that more than 40% of the temperature signal can be reconstructed at depth from remotely-sensed data using a simple statistical method (called synthetic fields) and that the complementary use of in situ measurements (denoted combined fields in Figure 2) improves the estimation by an additional 10-20%.



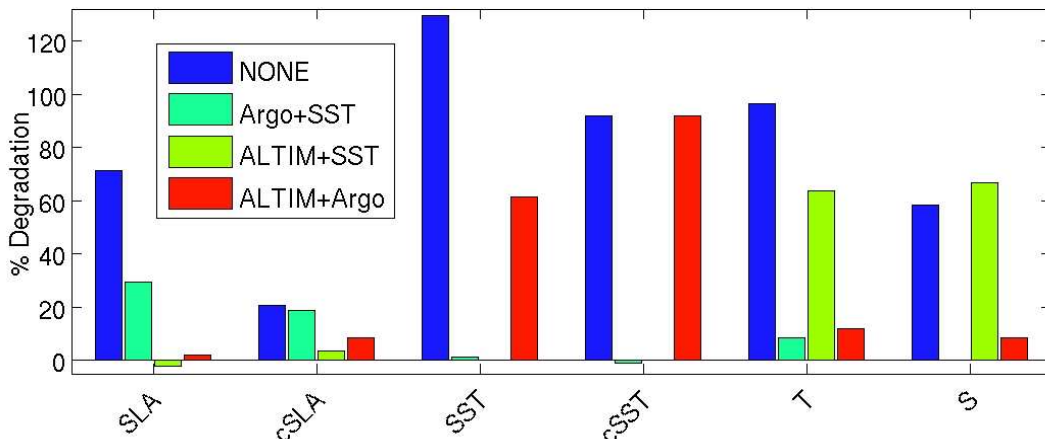
**Figure 2: RMS (solid lines) and mean (dotted lines) error in predicting sub-surface temperature (left) and S (right) anomalies using Levitus monthly mean climatology (red), synthetic fields (blue), combined fields (green); adapted from Larnicol et al. (2006).**

*b. Short-range Mesoscale Prediction Systems*

Through a series of OSEs, Oke and Schiller (2007) compare the relative impact of Argo, SST and SLA observations on an eddy-resolving ocean reanalysis. Briefly, they use a  $1/10^\circ$  resolution ocean model together with an ensemble OI (EnOI) system (Oke et al. 2008) to perform OSEs for a 6-month period at the start of 2006. In these experiments along-track altimeter, AMSR-E SST (AVHRR SST are withheld), Argo T and S and XBT data are assimilated. OSEs include a run with no data assimilation (NONE); a run that assimilates all observations referred to above (ALL); a run that assimilates Argo and SST data (Argo+SST), altimeter and SST data (ALTIM+SST), and altimeter and Argo data (ALTIM+Argo). The results from these OSEs are summarised in Figure 3, where a normalised measure of skill is presented for each OSE showing the % Degradation relative to a control run, following Benkiran et al. (2008):

$$\% \text{ Degradation} = 100 \times (\text{RMSE}_{\text{OSE}} - \text{RMSE}_{\text{REF}}) / \text{RMS}_{\text{OBS}}, \quad (1)$$

where RMSE is the root mean-squared error,  $\text{RMSE}_{\text{OSE}}$  is the RMSE for a given OSE,  $\text{RMSE}_{\text{REF}}$  is the RMSE for a reference model run and  $\text{RMS}_{\text{OBS}}$  is the RMS of observed fields. For Figure 3,  $\text{RMSE}_{\text{OSE}}$  is for each OSE, where a different data types are withheld, and  $\text{RMSE}_{\text{REF}}$  is for a model run where altimetry, AMSR-E SST, and Argo T/S data are assimilated. The results presented in Figure 3 include statistics for coastal SLA (cSLA) and coastal SST (cSST), where the SLA comparisons are restricted to coastal tide gauges and where the SST comparisons are restricted to locations where the water depth is less than 200 m. With the exception of cSLA, all RMSE estimates are based on comparisons with both assimilated and withheld observations. About half of the available Argo data are assimilated; AMSR-E data are only assimilated for 1 day in 7; all altimeter data are assimilated, but they are spatially and temporally averaged to form super-observations prior to assimilation. Figure 3 indicates that when no data are assimilated, the RMSE increases by between 20 and 130% of the observed signal. When altimeter data are withheld (see Argo+SST) the errors in SLA and cSLA increase by 30% and 20% respectively. When Argo data are withheld (see ALTIM+SST) the errors in T and S increase by 60-70% relative to the observed signal. When SST data are withheld (see ALTIM+Argo) the errors in SST, and particularly in coastal SST increase significantly, by almost 90% of the observed signal.

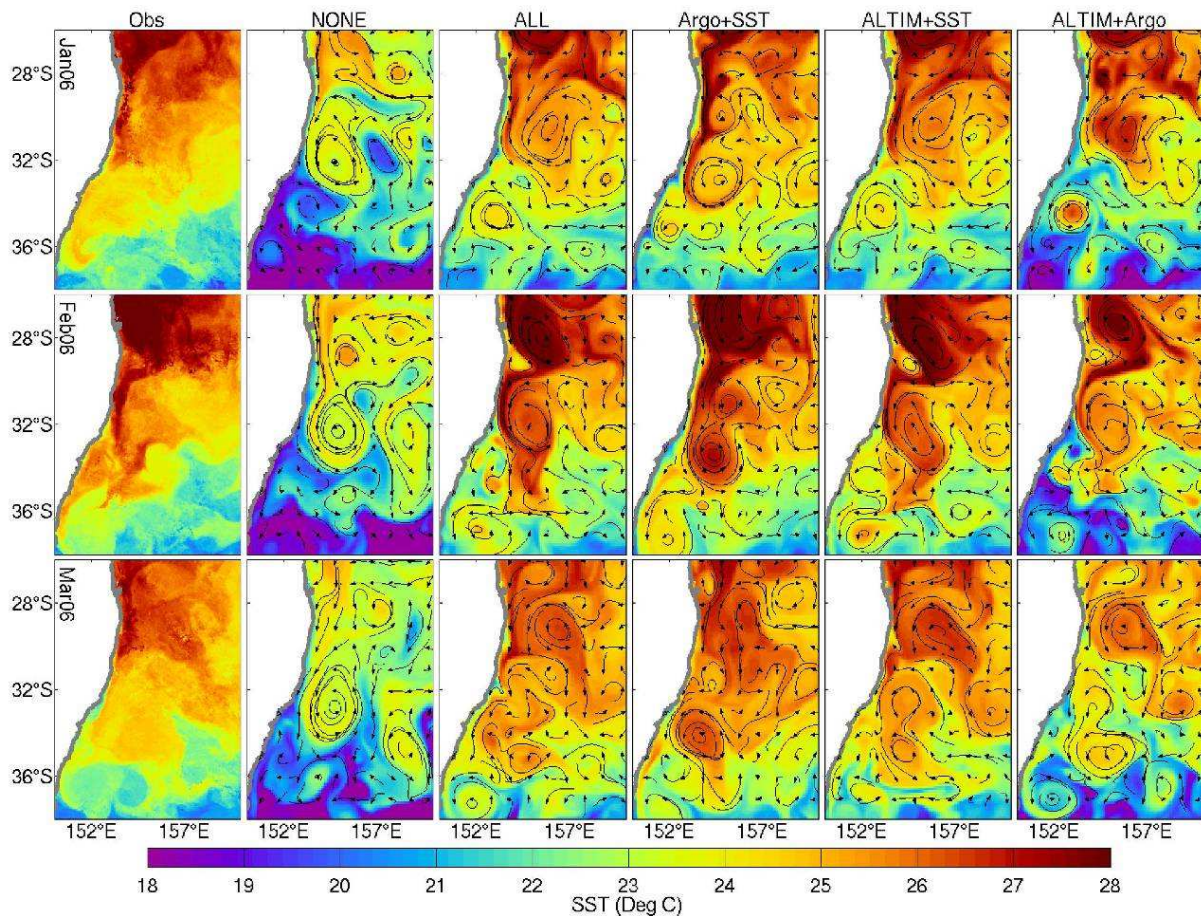


**Figure 3: The area-averaged % Degradation, from equation (1), for different variables. Coastal SLA (cSLA) is based on comparisons with withheld data from coastal tide gauges; and coastal SST (cSST) is based on comparisons with assimilated AMSR-E SST data where the bottom depth is less than 200 m; and comparisons with T and S are here restricted to the top 500 m. Area averages are computed for the region 90-180E and 60S-10N; adapted from Oke and Schiller (2007). When % Degradation is zero, the OSE performed as well as the reference run (see equation (1)); when % Degradation is 50% the RMSE increased by 50% of the observed signal compared to the reference run.**

A qualitative assessment of the OSEs presented by Oke and Schiller (2007) is presented in Figure 4. This shows a comparison between withheld AVHRR SST fields and modelled SST fields from each OSE. The fields in Figure 4 show that when SST is assimilated, the reanalysed SST is very similar to the observed SST. By contrast, when SST is not assimilated, the reanalysed SST differs significantly from observations. Consistent with the comparisons summarised in Figure 3, the non-assimilative case NONE produced the largest SST differences among the cases considered in Figure 4; these differences are reduced in ALTIM+Argo, but are smallest in the cases assimilating AMSR-E SST.

It is clear from Figure 4 that the mesoscale circulation is qualitatively similar for each OSE when altimetry is assimilated, but is quite different otherwise. For example, the position of the Tasman front in January and February, denoted by the separation of the EAC from the coast around 32S, is very similar in the OSEs that assimilate altimetry. Also note the robustness of the anti-cyclonic circulation around 152E and 37S in February and March; and the coastal meander near 29S in February that is evident in the observations and the OSEs that assimilate altimetry.

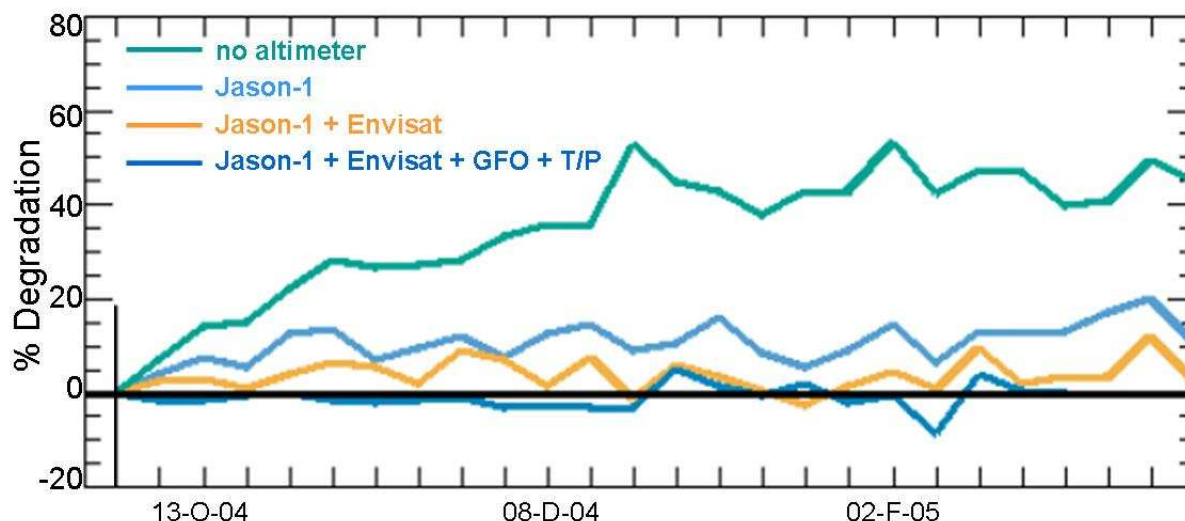
The main results from the study of Oke and Schiller (2007) are that each observation type brings complementary information to the GOOS. For example, satellite SST observations are the only observation type considered that constrains shallow seas and continental shelf circulation over wide shelves (see cSST in Figure 3); altimetry is the only observation type that properly constrains the mesoscale ocean circulation (Figure 3 and Figure 4); and Argo observations are the only observation type that constrains sub-surface T and S (Figure 3). These results indicate that while there is some redundancy for representing broad-scale circulation, all observation types are required for constraining mesoscale circulation models.



**Figure 4: Observed SST from (left) AVHRR (with-held data); and (columns 2-6) modelled SST and surface currents in the Tasman Sea for OSEs, as labelled in the title (NONE indicates no assimilation; ALL indicates Argo+SST+ ALTIM etc), for different time periods, as labelled in the left column, using the Bluelink reanalysis system; adapted from Oke and Schiller (2007).**

A series of OSEs has been conducted by Benkiran et al. (2008) to evaluate the impact of data from multiple altimeter missions on the forecast skill, over 7-days, using the Mercator Océan eddy resolving (5-7 km) forecasting system in the North Atlantic Ocean and the Mediterranean Sea. This system assimilates along-track altimeter data, SST and in situ profiles using a multivariate OI scheme. Specifically, Benkiran et al. (2008) sought to assess the degradation in the forecast skill when data from fewer altimeters were assimilated. They performed several 6-month simulations in which they assimilated all available SST and T/S profiles and altimeter data from 0 to 4 altimeters (T/P, Jason-1, Envisat and GFO). The OSEs were conducted during the tandem T/P and Jason-1 missions in 2004-2005 when data from 4 altimeters were available.

Figure 5 summarises the results of Benkiran et al. (2008), showing the % Degradation, from equation (1), where  $RMSE_{OSE}$  is for the OSEs that assimilate data from 0, 1, 2 and 4 altimeters; and  $RMSE_{REF}$  is for the run that assimilates data from 3 altimeters (Jason-1, Envisat and GFO). In this case, RMSEs are for SLA and are based on comparisons with Jason-1 observations. When no altimeter data are assimilated, there is effectively no predictive skill at the mesoscale. Conversely, Figure 5 suggests that skill is added to the Mercator system when data from 4 altimeters are assimilated, instead of just 3. These results are consistent with those presented by Pascual et al. (2006).



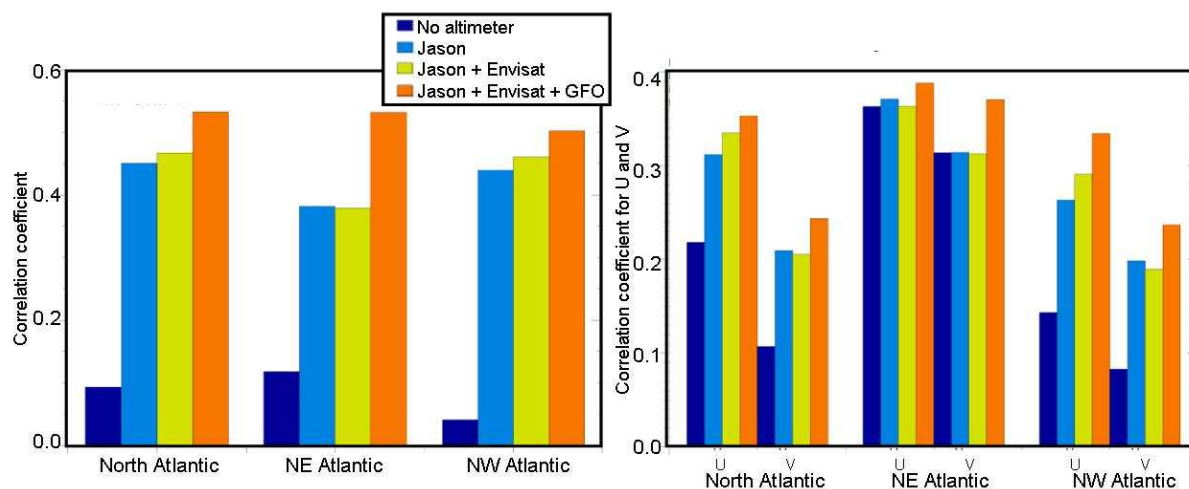
**Figure 5: A normalised measure of 7-day forecast error in the North Atlantic, when no altimeter data are assimilated and when data from 1, 2 and 4 altimeters are assimilated. Forecast skill is measured against the forecast error when Jason+Envisat+GFO data are assimilated (REF; see equation (1)). A positive % implies a degradation of the forecast skill, 0 is the baseline and negative means an improvement. RMSEs in equation (1) are based on comparisons with SLA observations from Jason-1; adapted from Benkiran et al. (2008).**

Benkiran et al. (2008) conducted their OSEs in a real-time context. That is, they performed OSEs to produce nowcasts under realistic conditions, excluding missing data due to latency of data availability. They produce 7-day forecasts that are initialised with each nowcast and also hindcasts, using all available data. Table 3 summarises their results, showing that if only SST and in situ T/S are assimilated (i.e., no altimetry) the error is large (up to ~13cm RMS). Table 3 also shows that with one altimeter, the situation improves for the hindcast (~8.5 cm), but the error is still large for the forecast (~10 cm RMS). Note also that to obtain error levels equivalent to the hindcast with only one altimeter, data from all 4 altimeters are needed for the initialisation of each 7-day forecast under realistic conditions; and data from at least 2 altimeters are required for the nowcast. These results are consistent with the conclusions of Pascual et al. (2008) who found that 4 altimeters in NRT is roughly equivalent to 2 altimeters in DT.

**Table 3: RMS of the difference between Jason-1 observation and 7-day forecast, Nowcast (real-time analysis) and hindcast (best analysis) for several OSEs where altimeter data from 0, 1, 2, 3 and 4 satellites are assimilated in addition to in situ T/S profiles and SST; adapted from Benkiran et al. (2008).**

SLA RMS difference	No altimetry	Jason-1 only	Jason-1 + Envisat	Jason-1+Envisat +GFO	Jason-1 +Envisat +GFO +T/P
7-day forecast (cm)		10.27	9.67	8.95	8.62
Nowcast (cm)		9.15	8.36	7.50	7.08
Hindcast (cm)	12.94	8.38	7.07	6.18	5.63

Results from a series of OSEs designed to assess the impact of different numbers of altimeters using the UK Met Office system are presented in Figure 6. Briefly, they use the  $1/9^\circ$  North Atlantic FOAM configuration together with an OI-based method of assimilation (Martin et al 2007) to run a series of three month integrations beginning in January 2006. The impact of different numbers of altimeters are assessed by comparing the modelled SLA with the assimilated along-track altimeter data, and comparing the modelled surface velocities with those derived from surface drifting buoys (which are not assimilated). These results quantify the improvements when one, two and three altimeters are added to the assimilated observations. The addition of the first altimeter seems to have the most impact. Interestingly, different satellites seem to give a different level of improvement. For example, the addition of GFO seems to result in a greater improvement than the addition of Envisat, although the order in which the different satellites are added may affect this result. The results are also different for different regions; surface velocities in the north-east Atlantic are better than surface velocities in the north-west Atlantic. This indicates that the mesoscale dynamics in the north-east are probably better constrained by the altimeters than in the north-west. The difference in the quality of surface winds in different regions probably also has a significant influence on the quality of the modelled surface velocities.



**Figure 6: Anomaly correlation between forecast (left) SLA and along-track altimetric SLA from all satellites and (right) forecast near-surface velocity and near-surface velocity derived from drifting buoys; based on a series of OSEs that assimilate SLA data from 0-3 satellites, using the  $1/9^\circ$  North Atlantic FOAM configuration for the first 3 months of 2006.**

### c. Seasonal Prediction Systems

The impact of the different components of the GOOS on ECMWF operational ocean analyses and seasonal forecasts is continuously assessed. In Vidard et al. (2007), the relative importance of the tropical in situ mooring arrays (TAO, TRITON, PIRATA), XBTs and the developing Argo float network was assessed, showing a major role for the moorings and in some ways a smaller contribution from Argo than might have been expected. There were three limitations in that paper: firstly Argo was assessed at a

relatively early stage of development (2001-2003), secondly the assimilation system did not use salinity data directly and thirdly altimeter data were not used.

In a later paper, Balmaseda et al. (2007) use the latest version of the ECMWF Ocean Re-Analysis System (ORA-S3) in which both salinity and altimeter data are assimilated. The ORA-S3 system is described by Balmaseda et al. (2008a). Briefly, ORA-S3 uses a global configuration of the HOPE model, with 1° resolution. A three-dimensional optimal interpolation (3DOI) scheme is used to assimilate T, S and SLA observations over a 10-day sequential update cycle. Argo and altimeter observations are systematically with-held from a series of 6-year reanalysis experiments (2001-2006). They show that Argo S is instrumental in correcting the salinity of the ECMWF analysis on a basin-wide scale. The effect of Argo temperature is noticeable in the Indian, Atlantic and Southern Oceans. In the Pacific, the impact is modest, being confined to the South American coast. The information content of Argo combines well with the altimeter information in most regions. Their results show that the assimilation of Argo observations results in improvements to seasonal forecasts of SST for most areas.

The relative contribution of Argo, altimeter and moorings to the skill of seasonal forecast is presented by Balmaseda et al. (2008b), based on a series of seasonal forecast experiments from ocean initial conditions generated by with-holding the specific observing system. The results demonstrate that Argo, altimeter and mooring observations contribute to the improvement of the skill of seasonal forecasts of SST. For example, Figure 7 shows the percentage reduction in the mean absolute error of 1-7 month SST forecast, averaged over the period 2001-2006. Figure 7 demonstrates that assimilation of Argo observations are particularly beneficial to SST forecasts in the eastern tropical Pacific (NINO4), altimeter data are particularly beneficial to the central Pacific and the north subtropical Atlantic (NINO3 and NSTRATL) and that mooring data have a significant positive impact on forecast skill across the entire tropical Pacific (NINO3 and NINO4). The positive impacts of Argo and mooring data on the forecast skill of SST in seasonal forecasts are also confirmed in JMA's system (Fujii et al. 2008a, see also the article on the seasonal forecast by Balmaseda et al. in this publication).

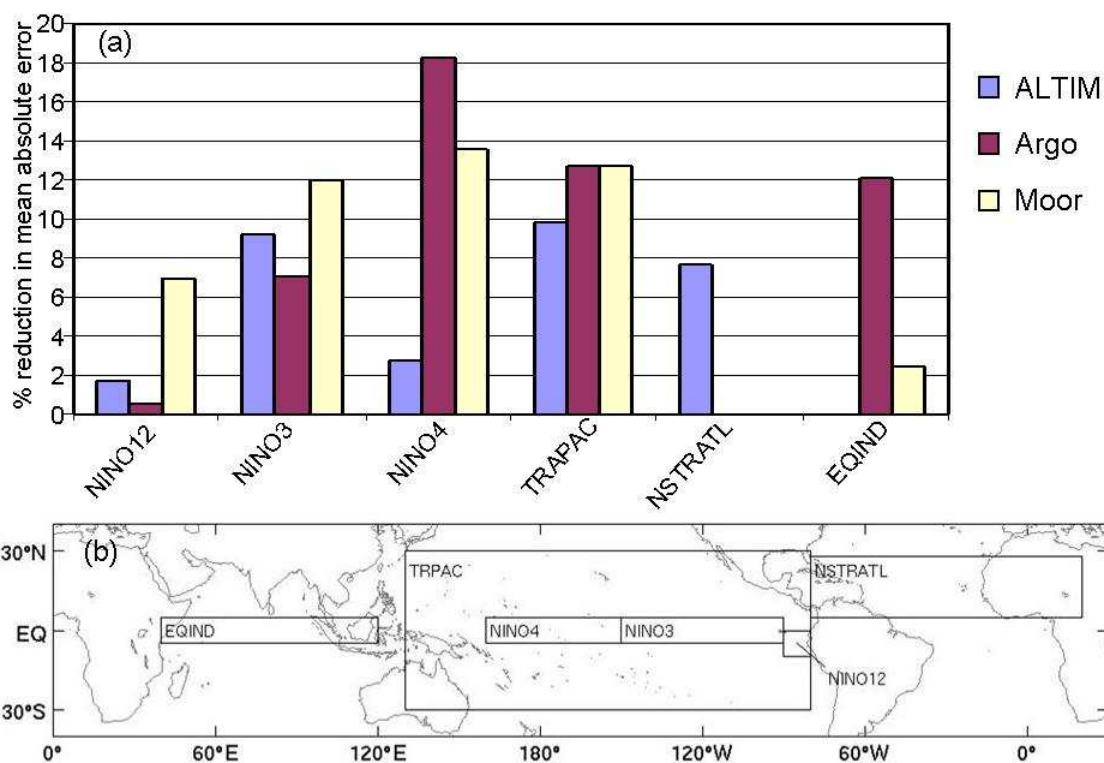
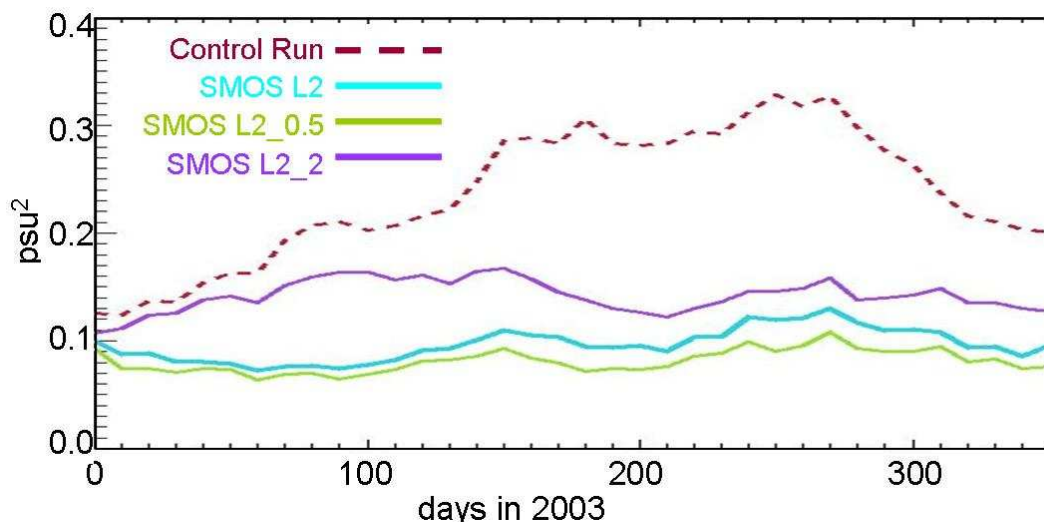


Figure 7: (a) Percentage reduction in the mean absolute error of 1-7 month SST forecasts for the period 2001-2006. ALTIM, Argo and Moor refer to the difference between the mean absolute error when altimeter, Argo and mooring data are withheld, respectively. Results are presented for the regions denoted in panel (b); adapted from Balmaseda et al. (2008b).

### 3. Observing System Simulation Experiments (OSSEs)

#### a. Assessing the potential impact of new/future observation types

The potential impact of new observation types on the forecast skill of the Mercator system has been assessed through a series of OSSEs. Tranchant et al. (2008) have conducted OSSEs to assess the potential impact of the remotely sensed salinity observations from SMOS (scheduled for launch in 2009), and Aquarius (planned for 2010). They have simulated various sets of SMOS and Aquarius datasets; and have subsequently assimilated these data in addition to real SST, T/S profiles and altimeter data in a version of Mercator Océan  $1/3^\circ$  North Atlantic forecasting system, using the multivariate SEEK filter. They show (Figure 8) that assimilation of remotely sensed salinity observations is likely to have a positive impact on the system performance by reducing the RMS error for salinity. They consider different scenarios, where they vary the magnitude of the SMOS observation error: the observation error for SMOS observed salinity is at the expected level for L2, half the expected level for L2\_0.5 and twice the expected level for L2\_2. The assumed observation errors for SMOS observations (SMOS L2) range from 0.2-2.5 psu, and vary in space and time. They are assumed to depend on factors such as brightness temperature, SST and wind speed. Based on these OSSEs, Tranchant et al. (2008) argue that the level of observation error will have a critical impact on the value of this new observation type to GODAE systems.

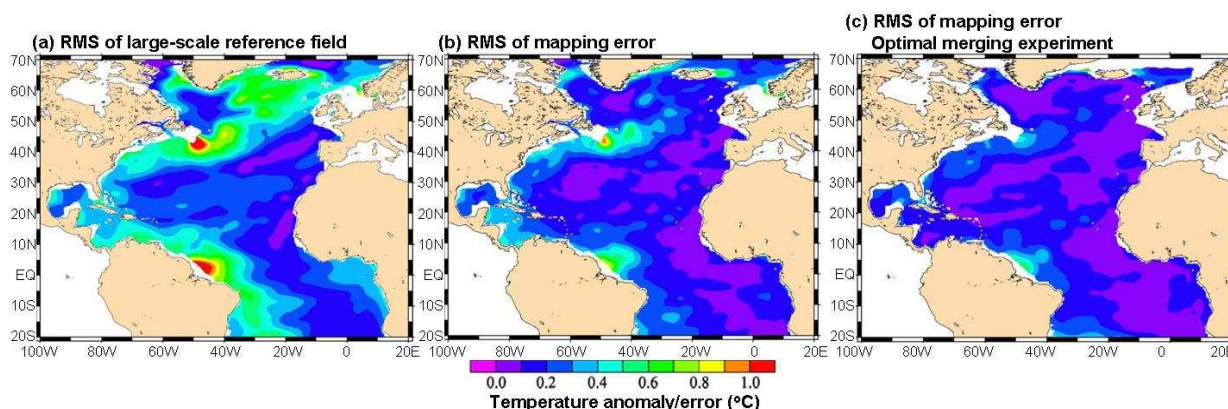


**Figure 8: Spatial average of the variance of the salinity difference between the “truth”, every ten days during 2003 for the overall domain, and three different estimates: control run (red dashed line), SMOS L2 (blue solid line), SMOS L2\_0.5 (green solid line, error divided by a factor of 2), SMOS L2\_2 (purple solid line, error multiplied by a factor of 2); adapted from Tranchant et al. 2008.**

More recently, Giraud et al (2008) have performed fraternal twin experiments using Mercator Océan systems, simulating surface velocity measurements in a manner that mimics that of synthetic aperture radar (SAR) data in coastal regions (see Chapron et al. 2005) or optical flow methods applied to SST or ocean colour data (see Vigan et al. 2000) in addition to more conventional observations (altimeter residuals, T/S profiles and SST). Their goal was to evaluate the potential of these innovative observations in the perspective of an eventual degradation of the altimeter coverage in the near future. They have used an eddy resolving forced simulation ( $1/12^\circ$  North Atlantic and Mediterranean NEMO configuration) to simulate the observations; and have subsequently assimilated them into the eddy permitting ( $1/3^\circ$ ) Mercator Océan forecasting system in the North Atlantic Ocean (Benkiran and Greiner 2007) following the fraternal twin experiment methodology. They conclude that ocean prediction systems could benefit from space-based velocity measurements - provided the observation errors remains below 7 cm/s; and provided they are used to complement satellite altimetry. Importantly, Giraud et al (2008) note that the impact of these surface velocity measurements should not be expected to compensate for a total loss of altimeter capacity.

### b. Large-scale variability

Based on a series of OSSEs, where “synthetic observations”, sampled from the CLIPPER model (Treguier et al., 1999) are mapped in the North Atlantic Ocean using an optimal interpolation method, the study by Guinehut et al. (2002) suggests that a 3° array of profiling floats cycling every 10 days, intended to represent an idealised Argo array, can retrieve a large proportion of the large-scale and low-frequency variability of the 200 m temperature fields (Figure 9b). Furthermore, Guinehut et al. (2004) show that by combining Argo-like data with synthetic T/S profiles projected from satellite SLA and SST, a much better estimate of the sub-surface temperature field can be reproduced (Figure 9c). The combination of the two data sets is particularly instrumental in reducing the aliasing due to the mesoscale variability. Consistent with the results presented in section 2, these results indicate that together, SLA, SST and in situ T/S observations can readily be mapped to realistically represent features of the mesoscale ocean circulation.



**Figure 9: (a) RMS of the large-scale and low-frequency reference field for temperature anomalies at 200 m depth; (b) the RMS of the mapping error for an idealised 3° resolution Argo array; and (c) the RMS of the mapping error of T at 200 m depth using an "optimal" combination of synthetic SLA, SST and T/S from an idealised 3° resolution Argo array; adapted from Guinehut et al. (2004).**

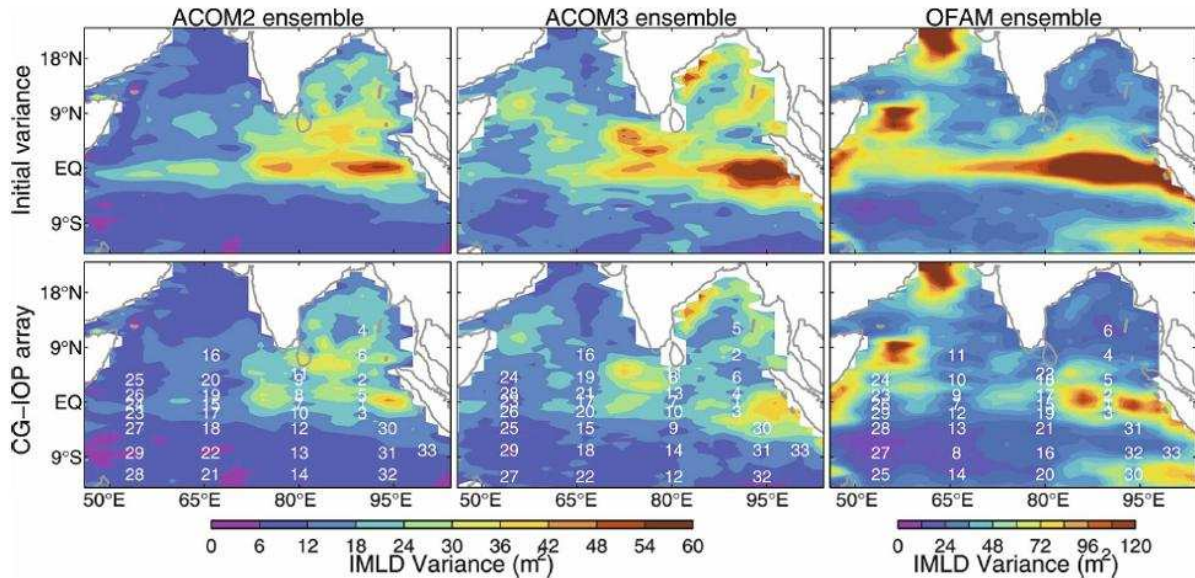
### c. Ensemble approach to Observing System Assessment and Design

Based on a series of OSSEs, using models of different resolution (coarse-intermediate-high resolution) that have different surface fluxes and are integrated for different periods, Sakov and Oke (2008) assess the CLIVAR-GOOS Indian Ocean Panel (CG-IOP) mooring array for the tropical Indian Ocean (Figure 10). Using an ensemble data assimilation technique they rank each of the proposed mooring sites (“1” is the “best”; “33” is the “worst”) in terms of reducing the analysis error variance. Additionally, Sakov and Oke (2008) apply their technique to objectively design an alternative mooring array for the tropical Indian Ocean and conclude that observations to the east and north-east of the region of interest, and observations due south of India are most important for resolving intraseasonal variability. They find that observations near 9°S, where seasonal Rossby waves dominate, are important for observing seasonal-to-interannual variability in the tropical Indian Ocean. These findings are consistent with many of the conclusions of Vecchi and Harrison (2007), Oke and Schiller (2007) and Ballabera-Poy et al. (2007), who also performed different flavours of OSSEs that were focussed on the tropical Indian Ocean.

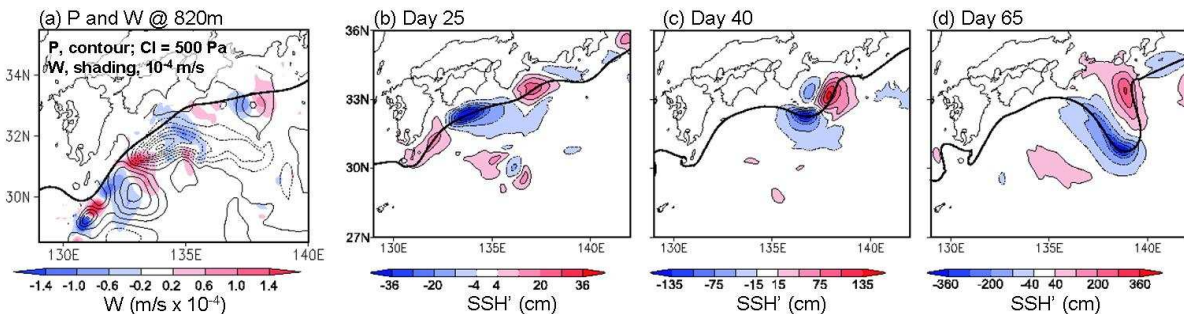
### d. Variational approach to Observing System Design

Calculation of singular vectors using a model's tangent linear approximation and its adjoint can be a very effective approach to observing system design. Leading singular vectors correspond to perturbations that lead to the fastest growing modes of a system. For example, Fujii et al. (2008b) compute the singular vectors for the MOVE system to investigate the types of perturbations that influence the large meanders in the Kuroshio Current. Specifically, they show that a perturbation that projects onto the leading singular vector around 133E and 31N leads to a growing perturbation and a further development of the large meander (Figure 11). An example of such perturbations is given in Figure 11a, showing perturbations to vertical velocity and pressure at 820 m depth. Fujii et al. (2008b) show that these perturbations cause cold

advection across the Kuroshio Current and downwelling to the north. This results in baroclinic instability and the development of an anticyclonic circulation in the deep layers. The corresponding anomalies to SSH that coincide with these developments are summarised in Figure 11b-d, showing the development of a large meander about 2 months after the initial perturbation. This analysis indicates that to properly predict the Kuroshio meander, a forecast model must be well constrained by data assimilation around 133E and 31N and particularly at depths of 1000 to 1500m. This implies that additional observations in that region are likely to benefit the forecast of the variability of the Kuroshio Current.



**Figure 10:** The modelled variance of the intraseasonal mixed layer from three different models (top) and the estimated analysis error variance (bottom) when observations from the CLIVAR/GOOS-Indian Ocean Panel mooring array are mapped using an ensemble approach. The numbers in each panel denote the mooring locations and the ranking of each location (i.e., the locations marked “1” are the “best” location); adapted from Sakov and Oke (2008).



**Figure 11:** (a) Perturbation fields for pressure (contour; dotted lines are negative) and vertical velocity (shading; positive is downward) at 820 m depth; and (b-d) SSH anomalies (scales are different for each panel) that result from the perturbations represented in panel (a) at day 0. Thick lines show the axis of Kuroshio Current in the background state; adapted from Fujii et al. (2008b).

#### 4. Emerging Techniques

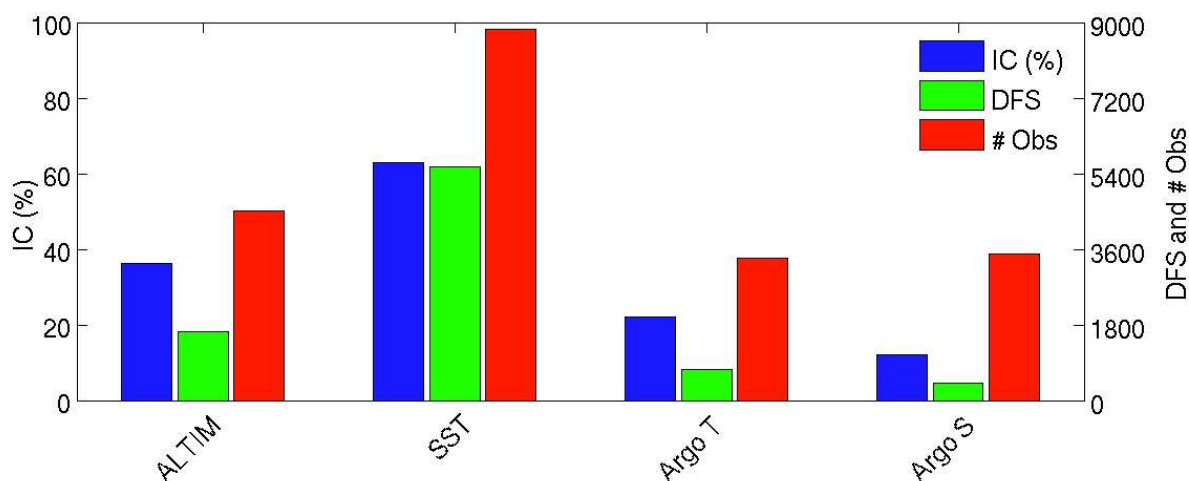
Within NWP, techniques have been developed to routinely evaluate observation value and understand patterns of observation impact on forecast skill (e.g., Rabier et al. 2008). The techniques are designed to quantify the information provided by observations in the assimilation and in the forecast. These diagnostics are important since all observations do not have equal value in reducing forecast error because of what is measured, where and when the measurements are taken, and the accuracy of the measurements. These new techniques can be broadly characterized in terms of either measuring the impact of observations on the analysis or on the forecast.

*a. Analysis Sensitivity Procedure*

Common to the application of most statistical data assimilation systems, such as OI, EnOI, EnKF and 3DVar is the employment of a gain matrix,  $\mathbf{K}$ . The gain matrix depends on the observation and background error covariances assumed by the system. A simple function of the gain matrix, referred to here as the influence matrix, is given by  $\mathbf{K}^T \mathbf{H}^T$  (Cardinali et al. 2004). The influence matrix is a well-known concept in statistical fields, but has only recently been derived for modern data assimilation systems. The influence matrix can diagnose the relative contributions of observations and background fields to a given analysis; the relative influence of different observation types; and the effective degrees of freedom of signal (DFS) in different data types. For example, Cardinali et al. (2004) report that for the ECMWF system, the global observation influence for each assimilation is around 15%. That means that 85% of each analysis is attributable to the background field. Of those observations, Cardinali et al. (2004) also report the relative contributions of different observation types for a typical analysis (e.g., 25% surface based and 75% satellite-derived observations).

Generally, analysis of the influence matrix can help identify low-influence and high-influence observations. We might expect low-influence observations to be those either in data-rich areas or those assumed to have high observation errors. Conversely, high-influence observations tend to be in data-sparse or dynamically active regions, and those assumed to have low observation errors. Importantly, analysis of the influence matrix does not necessarily quantify the value of observations – rather it quantifies how much of the observations are used by an assimilation system given the assumed error estimates therein.

Using the practical method of Chapnik et al. (2006) for approximating the diagonal elements of the influence matrix, known as the self-sensitivities, we present some preliminary estimates of the Information Content (IC) and the DFS of different observation types for the Bluelink system in Figure 12. Figure 12 shows the number of observations in the Australian region, divided by data type, where each observation counted may represent a super-observation of multiple measurements (see Oke et al. 2008). The DFS is the estimated sub-trace of  $\mathbf{K}^T \mathbf{H}^T$  for each data type and is an estimate of the effective number of independent observations assimilated. The IC, presented as a percentage, is based on the ratio of the DFS and the number of observations. Based on these results, it appears that both altimetry and SST observations are well used by the Bluelink system. However, information from the Argo data is clearly not extracted by the Bluelink system in an optimal way. This suggests that the error estimates used by the Bluelink system (Oke et al. 2008) need to be revisited and refined.



**Figure 12: Preliminary estimates of the Information Content (IC; %), degrees of freedom of signal (DFS) and the number of assimilated super-observations (# Obs) for the Bluelink reanalysis system in the region 90-180°E, 60°S-equator, computed for 1 January 2006. The scale for the IC is to the left and the scale for the DFS and # Obs is to the right.**

### *b. Adjoint-based Procedure*

The sensitivity of a forecast to observations can be computed using the adjoint of the assimilation procedure ( $\mathbf{K}^T$  where  $\mathbf{K}$  is the gain matrix) together with the adjoint of the forecast model (Langland and Baker, 2004). The procedure can be used to assess the impact of any, or all, of the observations used in the assimilation on reducing a measure of the forecast error. The forecast error cost function is typically formulated to represent errors in the position and strength of oceanographic features of interest (e.g., ocean fronts and eddies), although model prognostic variables can be directly used (e.g., T, S, velocity). One advantage of the adjoint-based method is that it is not necessary to selectively add or remove observations or observing systems from the assimilation, as in conventional OSEs, which change the analysis constraints on the remaining data and can alter the outcome of the assimilation and subsequent forecast. Another advantage of the method is that observation impacts can be partitioned for any subset of the data assimilated; by instrument type, observed variable, geographic region, or vertical level. Also, the method is computationally inexpensive, compared to OSEs, and can be readily integrated into an operational ocean analysis/forecast system. It is also worth noting however, that this technique requires a linear assumption that is probably most appropriate for short-term (days) forecast problems, but may not be valid for longer term (months) forecast problems, such as seasonal prediction using a coupled ocean-atmosphere model.

Since ocean observing and assimilation/forecast systems are in continuous evolution, an efficient procedure that allows the impact of observations to be regularly assessed is clearly needed. In addition to monitoring the impact of observations on the forecast, the adjoint-based procedure can be used to develop more optimal observing networks. Forecast error sensitivities to hypothetical observations in combination with the real observations can easily be calculated. Extreme beneficial impacts from the hypothetical observations will indicate a need for increased routine monitoring at those locations. Finally, the adjoint-based procedure is useful for adaptive sampling. In this application, the impact of adding targeted observations to the regular observing networks on reducing forecast error at some future time over a specific region is assessed. Knowledge of forecast error sensitivities in advance can be used by operators to target deployments of observations, such as ocean gliders, to areas where the potential for forecast error growth is high. The impact of the additional glider observations on the actual forecast error can then be directly assessed in the next assimilation update cycle. An adjoint-based data impact system is being developed for the global HYCOM forecast system in the U.S. The system is expected to be producing routine assessments of the GOOS by 2010.

## **5. Conclusions**

One of the purposes of this paper is to identify the robust and important results from a series of OSEs and OSSEs performed by different research teams for a range of different applications. One recurring result from different OSEs includes the apparent complimentary nature of different observation types (e.g., Guinehut et al. 2004, Figure 1 and Figure 9; Larnicol et al. 2006, Figure 2; Oke and Schiller 2007, Figure 3, Figure 4; Balmaseda et al. 2008b, Figure 7). This means that none of the observation types in the GOOS is redundant. Each different observation type brings unique contributions to the GOOS and all observation types should be routinely assimilated by forecast and reanalysis products; and more importantly maintained by the international community.

Another result that is common to many studies is the necessity of assimilation of altimeter data to represent mesoscale variability (e.g., Oke and Schiller 2007, Figure 3, Figure 4; Martin et al. 2007, Figure 6; Pascual et al. 2008, Table 1 and

Table 2; Benkarin et al. 2008, Figure 5, Table 3). Moreover, a couple of studies demonstrated that for NRT applications data from 4 altimeters is needed to obtain errors that are comparable to systems using 2 altimeters in delayed-mode (e.g., Pascual et al. 2008, Table 1 and

Table 2; Benkiran et al. 2008, Table 3). This is because of differences in the quality of NRT data compared to DT data; and due to the latency of data delivery in NRT. This has important implications for

operational systems that use altimeter data in NRT. In the event that the altimeter coverage is compromised in the future, the study by Girard et al. (2008) indicates that assimilation of unconventional data types, like surface velocities derived from SAR data or other remote sensing products, may go some way towards filling the void. However, they note that these data types will not compensate for a total loss of altimeter observations. More optimistically, it is anticipated that future wide-swath altimeter missions will go some way towards solving this problem, with more complete global coverage and shorter repeat cycles.

Several studies have demonstrated the importance of Argo observations. These include several OSEs and OSSEs based on analysis systems (e.g., Guinehut et al. 2004, Figure 9; Larnicol et al. 2006, Figure 2) and OSEs based on both short-range and seasonal prediction systems (e.g., Oke and Schiller 2007, Figure 3, Figure 4; Balmaseda et al. 2008b, Figure 7; Fujii et al. 2008b). Several of these studies specifically noted that Argo is the only observation platform that provides global-scale information for constraining salinity. We reiterate, however, that the ocean forecasting community must continue to improve the prediction systems to more effectively use in situ observations including Argo profiles, as suggested in section 4a.

All GODAE systems considered in this paper include SST observations as an essential core data set. The consistent uptake of SST observations is a credit to the GHRSSST program that provides high level quality controlled SST data in NRT. Based on the OSEs presented by Oke and Schiller (2007; Figure 3, Figure 4), it is clear that SST is also essential for representing mesoscale ocean variability, particularly in shallow regions where altimeter data becomes somewhat unreliable.

The versatility of OSSEs and variational data assimilation techniques are also demonstrated in this paper, where it is shown that insight into observing strategies for resolving specific processes, like the Kuroshio meander (Fujii et al. 2008b, Figure 11), and specific time-scales of variability (e.g., Sakov and Oke 2008, Figure 10) can be gained. The impact of new observation types, like surface salinity observations, has also been assessed, with promising results (Tranchant et al. 2008, Figure 8).

We note that many groups from the NWP community routinely provide statistics on data impacts; in some cases - every day for every assimilation cycle. The methods discussed in sections 2 and 3 of this paper (OSE and OSSEs) are very expensive - and as a result are not applied routinely. They are also of limited value. For example, they will not automatically identify the impacts of changes in the Argo array - as the total number of Argo floats fluctuates and their spatial distribution changes. By contrast, as the NWP community have demonstrated, the routine application of computationally efficient methods, such as those referred to in section 4 can readily be applied to operational systems in NRT - and can potentially support the maintenance and development of the GOOS on an ongoing basis.

During the new sustained phase of GODAE, so called GODAE OceanView, more efforts should be given to OSE and OSSE studies; and to the performance of observing system evaluation in NRT using methods like those described in section 4. Development of international coordination and cooperation on these topics is essential. This is needed to provide consistent and educated responses to space and in-situ agencies and organisations in charge of a sustained global and regional ocean observing systems. A joint GODAE OceanView/OOPC Observing System Evaluation Task Team has been proposed with these objectives in mind.

#### **Acknowledgement**

The participants of the OOPC-GODAE meeting on OSSEs and OSEs, UNESCO/IOC, Paris, France in November 2007 are gratefully acknowledged for their contributions to this work. The following organisations and agencies are acknowledged for their financial support: Bureau of Meteorology; CSIRO; European Space Agency; French Service Hydrographique Océanographique de la Marine (SHOM); Japan Meteorological Agency; the Japanese Ministry of Education, Culture, Sports, Science and Technology; Office of Naval Research; and the Royal Australian Navy. Satellite altimetry is provided by NASA, CNES and NOAA and SST observations are provided by NOAA. Argo data are provided by the Coriolis and USGODAE data centres.

#### **References**

Balmaseda, M. A., D. Anderson, and A. Vidard, (2007) Impact of Argo on analyses of the global ocean. *Geophysical Research Letters*, **34**, L16605, doi:10.1029/2007GL030452.

- Balmaseda, M. A., Arthur Vidard and David Anderson (2008a) The ECMWF ORA-S3S ocean analysis system. *Monthly Weather Review*, **136**, 3018-3034.
- Balmaseda, M. A., and D. Anderson (2008b) Impact on initialization strategies and observations on seasonal forecast skill. *Geophysical Research Letters*, submitted.
- Ballabrera-Poy, J., E. Hackert, R. Murtugudde, and A. J. Busalacchi, (2007) An observing system simulation experiment for an optimal moored instrument array in the tropical Indian Ocean. *Journal of Climate*, **20**: 3249-3268.
- Benkiran, M and E. Greiner (2007) Data assimilation of drifter velocities in the Mercator Océan system. *Mercator Océan quarterly newsletter*, **25**, [www.mercator-cean.fr/documents/lettre.lettre\\_25\\_en.pdf](http://www.mercator-cean.fr/documents/lettre.lettre_25_en.pdf).
- Benkiran, M., E. Greiner, S. Giraud St Albin, E. Dombrowsky, D. Jourdan and M. Faillot (2008) Impact study of the number Space Altimetry observing systems on the altimeter data assimilation in the Mercator-Ocean system. In preparation.
- Brachet S., P. Y. Le Traon, C. Le Provost (2004), Mesoscale variability from a high-resolution model and from altimeter data in the North Atlantic Ocean, *Journal of Geophysical Research*, **109**, C12025, doi:10.1029/2004JC002360.
- Cardinala, C., S. Pezzulli and E. Andersson (2004) Influence-matrix diagnostic of a data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, **130**, 2767-2786.
- Chapnik, B., G. Desroziers, F. Rabier and O. Talagrand (2006) Diagnosis and tuning of observational error in a quasi-operational data assimilation setting, *Quarterly Journal of the Royal Meteorological Society*, **132**, 543-565.
- Chapron, B. F. Collard and F. Arduin (2005) Direct measurements of ocean surface velocity from space: interpretation and validation. *Journal of Geophysical Research*, **110**, C07008, doi 10.1029/2004JC002809.
- Ducet, N., P.-Y. Le Traon, and G. Reverdin, 2000: Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2. *J. Geophysical Research*, **105**, 19 477–19 498.
- Fujii, Y., T. Yasuda, S. Matsumoto, M. Kamachi, and K. Ando (2008a) Observing System Evaluation (OSE) using the El Niño forecasting system in Japan Meteorological Agency. Proceedings of the oceanographic society of Japan fall meeting (in Japanese).
- Fujii, Y., H. Tsujino, N. Usui, H. Nakano, and M. Kamachi (2008b) Application of singular vector analysis to the Kuroshio large meander. *Journal of Geophysical Research*, **113**, C07026, doi:10.1029/2007JC004476.
- Giraud St Albin, S., M. Benkiran, E. Greiner, S. Vrac (2008). De l'utilité des observations de vitesse de surface pour l'assimilation de données dans les systèmes opérationnels de prévision océanique. *Revue de l'électricité et de l'électronique published by the Société de l'Electricité, de l'Electronique et des Technologies de l'Information et de la Communication (SEE)*, Paris, France, ISSN 1265-6534, N°2, Pages 63-75 (in french)
- Guinehut, S., G. Larnicol and P. Y. le Traon (2002) Design of an array of profiling floats in the North Atlantic from model simulations. *Journal of Marine Systems*, **35**, 1-9.
- Guinehut, S., P. Y. Le Traon, G. Larnicol and S. Phillips (2004) Combining Argo and remote-sensing data to estimate the ocean three-dimensional temperature fields – a first approach based on simulated observations. *Journal of Marine Systems*, **46**, 85-98.
- Langland and Baker (2004) Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189-201.
- Larnicol, G., S. Guinehut, M.-H. Rio, M. Drevillon, Y. Faugère and G. Nicolas (2006) The Global Observed Ocean Products of the French Mercator Project, *Proceedings of 15 Years of progress in radar altimetry Symposium, ESA Special Publication*, pp. 614.
- Le Traon, P.Y. and G. Dibarboure (2002) Velocity mapping capabilities of present and future altimeter missions: The role of high frequency signals. *Journal of Atmospheric and Oceanic Technology*, **19**, 2077-2088.
- Martin, M.J., A. Hines, M. J. Bell (2007) Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact. *Quarterly Journal of the Royal Meteorological Society*, **133**, 981–995.
- Oke, P. R., and A. Schiller (2007) A model-based assessment and design of a tropical Indian Ocean mooring array, *Journal of Climate*, **20**: 3269-3283.
- Oke, P. R., and A. Schiller (2007) Impact of Argo, SST and altimeter data on an eddy-resolving ocean reanalysis. *Geophysical Research Letters*, **34**, L19601, doi:10.1029/2007GL031549.
- Oke, P. R., G. B. Brassington, D. A. Griffin and A. Schiller (2008) The Blueink Ocean Data Assimilation System (BODAS), *Ocean Modelling*, **21**: 46-70.
- Pascual A., Y. Faugère, G. Larnicol and P.-Y. Le Traon (2006) Improved description of the ocean mesoscale variability by combining four satellite altimeters, *Geophysical Research Letters*, **33**, doi:10.1029/2005GL024633.

- Pascual A., M. I. Pujol, G. Larnicol, P. Y. Le Traon, M. H. Rio (2007) Mesoscale Mapping Capabilities of Multisatellite Altimeter Missions: First Results with Real Data in the Mediterranean Sea. *Journal of Marine Systems*, **65**, 190-211.
- Pascual A., C. Boone, G Larnicol, P.Y. Le Traon (2008) On the quality of real time altimeter gridded fields: comparison with in situ data, *Journal of Atmospheric and Oceanic Technology*, in press.
- Rabier, F., P. Gauthier, C. Cardinali, R. Langland, M. Tsyulnikov, A. Lorenc, P. Steinle, R. Gelaro, and K. Koizumi (2008) An update on THORPEX-related research in data assimilation and observing strategies. *Nonlinear Processes in Geophysics*, **15**, 81-94.
- Sakov, P., and P. R. Oke (2008) Objective array design: Application to the tropical Indian Ocean, *Journal of Atmospheric and Oceanic Technology*, **25**: 794-807.
- Tranchant, B., C.E. Testut, L. Renault, N. Ferry, F. Birol, P. Brasseur (2008) Expected impact of the future SMOS and Aquarius ocean surface salinity missions in the Mercator Ocean operational systems: new perspectives to monitor ocean circulation. *Remote Sensing of Environment*, **112**, 1476-1487.
- Treguier A.-M., Coauthors, 1999: The CLIPPER project: High resolution modelling of the Atlantic. *International WOCE Newsletter*, No. 36, WOCE International Project Office, Southampton, United Kingdom, 3-5.
- Vecchi, G. A., and M. J. Harrison, 2007: An observing system simulation experiment for the Indian Ocean. *Journal of Climate*, **20**, 3300-3343.
- Vidard, A., D. L. T. Anderson, and M. Balmaseda (2007) Impact of ocean observation systems on ocean analysis and seasonal forecasts, *Monthly Weather Review*, **135**, 409-429.
- Vigan, X. et al. (2000). Sea surface velocities from the sea surface temperature image sequence, 1. Method and validation using primitive equation model output. *Journal of Geophysical Research*, **105**, 19499-19514.