Observing System Design and Assessment

Peter R. Oke and Terence J. O'Kane

CAWCR, CSIRO Marine and Atmospheric Research and Wealth from Oceans National Research Flagship

Abstract

The use of models and data assimilation tools to aid the design and assessment of ocean observing systems is increasing. The most commonly used technique for evaluating the relative importance of existing observations is Observing System Experiments (OSEs), and Observing System Simulation Experiments (OSSEs). OSEs are useful for looking back, to evaluate the relative importance of existing of past observational components, while OSSEs are useful for looking forward, to evaluate the potential impact of future observational components. Other methods are useful for looking at the present, and are therefore most useful for adaptive sampling programs. These include analysis self-sensitivities, and a range of ensemble-based and adjoint-based techniques, including breeding, adjoint sensitivity, and singular vectors. In this chapter, the concepts for observing system design and assessment are introduced. A variety of different methods are then described, including examples of oceanographic applications of each method.

Introduction

The use of models and data assimilation tools to aid the design of observing systems has a long history in numerical weather prediction (NWP; e.g., Kuo et al., 1998; Bishop et al. 2001; 2003) and is gaining momentum in the ocean modelling community (e.g., Oke et al. 2009). Methods for observing system design and assessment range from basic analysis of models to assess de-correlation length- and time-scales, signal-to-noise ratios, and covariance of different variables and different regions. Classical model-based approaches to observing system design and assessment involve observing system simulation experiments (OSSEs) and observing system experiments (OSEs). More sophisticated methods have emerged as a result of advances in data assimilation methodology, and there are now a suite of ensemble-based and adjoint-based techniques for designing observing systems and evaluating the impact of observations on assimilating models.

Observing system design and assessment has a long history in NWP. Most NWP applications relate to adaptive sampling. Adaptive sampling is the problem of identifying where additional observations should be made to better initialise a forecast. Typical examples of adaptive sampling programs in NWP relate to the prediction of extreme weather, like hurricanes (e.g., Gelaro et al. 1999). The idea is that if additional observations are made where an instability is developing, or is likely to develop, then those observations can be used to better initialise an NWP forecast and therefore improve the skill of that forecast. Adaptive sampling in NWP arguably began in 1947, when the hurricane reconnaissance program was established to observe of the location and intensity of hurricanes. In 1982, NOAA's Hurricane Research Division began research flights around hurricanes to improve the initialisation of NWP forecasts. They found that the error in the forecast tracks of hurricanes reduced by 25% as a direct result of their adaptive sampling program. In 2003, the World Meteorological Orginisation (WMO) initiated a program called THe Observing system Research and Predictability EXperiment (THORPEX). THORPEX was established with the intent to improve the accuracy of NWP forecasts of high-impact weather. Within THORPEX, the data assimilation and observation strategy working group was established to assess the impact of observations and various targeting methods to provide guidance for observation campaigns and for the configuration of the global observing system. For an excellent summary of THORPEX activities and results, the reader is referred to Rabier et al. (2008).

Ocean data assimilation capabilities have progressed rapidly since the beginning of the Global Ocean Data Assimilation Experiment (GODAE; www.godae.org/). A suite of analysis and forecast systems are now used routinely for operational and research applications. All GODAE forecast and analysis systems are underpinned by the Global Ocean Observing System (GOOS; www.ioc-goos.org) that is comprised of satellite altimetry, satellite sea surface temperature (SST) programs, delivered through the GODAE High Resolution SST effort (GHRSST; www.ghrsst-pp.org), and in situ measurements from the Argo program (Argo Science Team 1998), the tropical moored buoy (McPhaden et al. 1998), surface drifting buoys (www.aoml.noaa.gov/phod/dac), expendable bathythermograph (XBT; www.jcommops.org/soopip/; www.hrx.ucsd.edu) and tide gauge networks. Each of these observation programs are expensive and require a significant international effort to implement, maintain, process, and disseminate. Careful design and assessment of the GOOS is therefore warranted.

Observing system design and assessment activities in the oceanographic community are becoming more common. One of the key challenges for the oceanographic community is to adequately combine the efforts of researchers operating the climate domain, under CLIVAR (www.clivar.com; Heimbach et al. 2010), and those operating in the short-term forecasting domain, under GODAE (www.godae.org/ OSSE-OSE-home.html; Oke et al. 2009; 2010). CLIVAR activities tend to focus on climate monitoring and ocean state estimation, while GODAE activities tend to focus on mesoscale variability and short-range forecasting. Observational requirements for these different applications are likely to be quite different.

In this chapter, the concepts of observing system design and assessment are introduced, followed by a description of commonly used methods. The description of each method is intended to be practical, with less focus on theory and more focus on how things are actually done. For each method that is discussed, an oceanographic example is included, where possible. The chapter concludes with a short summary.

Concepts for observing system design and assessment

Before undertaking any activity that relates to observing system design and assessment, there are several key questions that need to be addressed. These questions relate to the motivation for establishing an observing system, practical limitations, and how the observations will be used.

The motivation for establishing an observing system is obviously important. What is it that the observing system is intended to monitor? This might be, for example, the heat content in a specific region, the volume transport of a current system, the variability of the thermocline depth, and so on. An observing system that is optimised to monitor a specific aspect of the ocean circulation is unlikely to be optimal for monitoring all other aspects of the circulation. For example, an observing system that is optimised for initialising a seasonal forecast system that seeks to predict the onset of El Nino will resolve dynamical features that vary on time-scales of El Nino like tropical instability waves, and is likely to be quite different to an observing system that is optimised to constrain an eddy-resolving ocean model that will resolve dynamical features that vary on shorter time-scales. So, the motivation for the observing system should be clear, and where the intended use of the observing system is broad, the optimisation strategy should attempt to reflect this as much as possible.

An understanding of what observations are feasible is important. This is likely to be dictated by budget, technology, and convenience. Deployment and maintenance of observations is usually expensive, so a well-design array that is easily deployed and maintained (e.g., with moorings along shipping lanes) may be essential. The budget may provide guidance on the number and types of instruments that can be considered (e.g., number and type of moorings, gliders, Argo floats, drifting buoys, etc). Many studies begin with a specification that, for example, the observation array may consist of up to 10 moorings that each measure temperature and velocity between the surface and 300 m depth; and ask the question, where should those moorings be deployed?

The question of how the observations will be used is difficult because in most cases there are likely to be multiple users, each processing the observations using different methods. For example, observations might be assimilated into a number of models using different assimilation methods; or observations might be gridded using a variety of techniques. It is typical, to assume that a specific analysis or assimilation system will be used to objectively map the observations. In this case, it is important to be clear about the characteristics and limitations of the particular analysis tool of choice. A better approach is to use a multi-system (e.g., multi-model) approach, where several systems are used to evaluate different observation arrays. This is the aspiration of many of the activities under GODAE OceanView (see www.godae.org/ OSSE-OSE-home.html).

The density of observations required to monitor a given process is largely dictated by the de-correlation length-scales of the fields that are to be observed. This characteristic determines how far apart observations can be made before important features are missed. Similarly, de-correlation time-scales determine how frequency observations should be made. The use of models to determine length- and time-scales is often fraught with difficulty, because of sub-grid-scale parameterisations within models largely determine these scales, and those parameterisations are generally inaccurate and are sensitive to many subjective choices made by the model developers (e.g., O'Kane and Frederiksen 2008a; Frederiksen and O'Kane 2008).

Some more subtle characteristics also become important for the design of observing systems. The co-variability of the ocean is critical. Are there locations or quantities that are particularly indicative of the entire system that is to be observed? That is, is there a specific location that is the *pulse* of the region of interest? The Southern Oscillation Index (SOI) is a good example of this. The SOI is calculated from variations in the air pressure difference between Tahiti and Darwin. Periods of sustained negative SOI usually correspond to El Nino events that are characterised by warming in the central tropical Pacific Ocean, a decrease in the trade winds, and reduced rainfall over much of Australia.

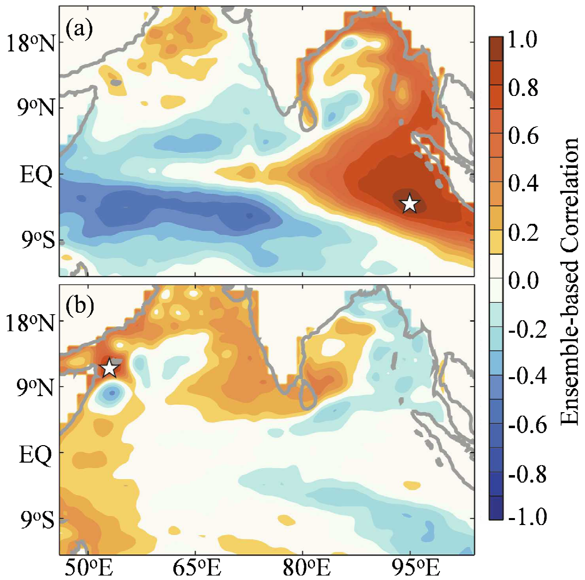

Figure 1: Examples of the ensemble-based correlation between sea-level at a reference location, denoted by the star, and sea-level in the surrounding region. Adapted from Sakov and Oke (2008).

An example of how a model can be used to identify the *pulse* of the ocean is presented in Figure 1, showing two examples of correlation fields from an ensemble-based data assimilation system (Sakov and Oke 2008). Ensemble-based assimilation systems use an ensemble of anomalies (also called perturbations or modes) to implicitly represent the system's background error covariance. The background error covariance determines how an observation-model difference is projected onto the model state during the assimilation step. So the ensemble-based correlation (or covariance) between an observable variable at reference location and the rest of the model state represents the effective *foot-print* of an observation at that reference location. The examples presented in Figure 1shows the ensemble-based correlation between sea-level at different reference locations and sea-level in the surrounding region. The regions where the amplitudes of these correlations are large correspond to regions where an observation from that reference location will have a significant impact.

The first example, shown in Figure 1a, indicates that an observation in the eastern Indian Ocean, off Java, is well correlated with sea level along the coast and over a very broad region. The spatial structure of the correlation map shows a dipole structure. This structure is observed in several previous studies (Chambers et al. 1999; Feng et al. 2001; Wijffels and Meyers 2004; Rao and Behera 2005). Also, the footprint of the positively correlated region reflects Rossby–Kelvin wave patterns. This indicates that observations offshore of Indonesia are likely to be particularly useful for constraining a data assimilating model that uses an ensemble like that described by Sakov and Oke (2008).

The second example, shown in Figure 1b, indicates that sea level off Somalia is relatively uncorrelated with sea level across the tropical Indian Ocean. The region off Somalia is dominated by mesoscale variability that spawns from the energetic and highly variable boundary currents in this region. While the mesoscale variability in this region is well organized (Schott and McCreary 2001), its variability is apparently somewhat chaotic and is characterized by short de-correlation length scales. This suggests that, while many observations may be required in the north-west tropical Indian Ocean to adequately represent the variability there, an observation in this region will not impose a significant constraint on a data assimilating model that uses the ensemble described by Sakov and Oke (2008).

Like any optimisation problem, observing system design and assessment ultimately involves the quantification of how good an observing system is. Consequently, the most important question for any observing system design or assessment activity is: what is it we seek to minimise? This is quantified by a cost function, metric, or diagnostic. The possible metrics that could be minimised by are virtually unlimited. We might seek to minimise the analysis error variance of some quantity (e.g., temperature, salinity, velocity, thermocline depth) for some region (e.g., tropical Pacific Ocean, North Atlantic, etc.). We might seek to minimise the forecast error of some quantity in a given region. Or perhaps we seek to minimise uncertainty of an integrated quantity, such as the transport through a strait. We may even wish to minimise several quantities (e.g., temperature and velocity error), which may require some sort of normalization, or weighting, that reflects the variance of different variables or relative importance for a given application. In every case, we must define a cost function, or metric, that we seek to minimise. The results will often depend heavily on this cost function (e.g., Sakov and Oke 2008).

Methods and Examples

Commonly used techniques for evaluating the benefits of different observation types and arrays include Observing System Experiments (OSEs), Observing System Simulation Experiments (OSSEs), analysis self-sensitivities, ensemble-based methods, and adjoint-based methods. All of these methods require some form of data assimilation. Of these methods, OSEs, OSSEs and analysis self-sensitivities can all be applied regardless of the assimilation technique used. By contrast ensemble-based and adjoint-based methods require specific tools for their application. Details of all of these methods, including examples, follow. Other methods, not described in detail here, that have also been applied to observing system design and assessment include genetic algorithms (Gallagher et al. 1991). Applications of genetic algorithms to oceanic applications include the optimisation of surface drifter deployments (Hernandez et al. 1995), and acoustic tomography arrays (Barth 1992).

Observing system Experiments - OSEs

The most commonly used method for employing assimilating models to assess observing systems is OSEs. OSEs generally involve the systematic denial, or withholding, of different observation types from a data assimilating model in order to assess the degradation in quality of a forecast or analysis when that observation type is not used. 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. It is therefore instructive to consider 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.

Results from OSEs can sometimes be difficult to interpret. Suppose four different observation types, from different platforms (e.g., Argo floats, satellite SST, altimetry, moorings) are typically assimilated. We might expect that there is some redundancy between these data types. For example, some of the information contained in an Argo profile is represented by altimetry (e.g., Guinhut et al. 2004). Similarly, some data in SST fields is also measured by Argo floats. If we withhold Argo data from an OSE, we might expect altimetry and SST to become more important, so the true value of an observation, or observation type, is difficult to really assess with OSEs.

In some cases, subtle details of the model/assimilation system can complicate the interpretation of OSEs. For example, Vidard et al. (2007) report a case when they with-held observations in the tropics. They found that with-holding this data degraded the circulation at high latitudes. This was puzzling. They traced this link back to the quality control system of their assimilation. An important step for any quality control system is a comparison with the model's background field. If observations differ significantly from the background field, they may be flagged as *bad*, and automatically with-held from assimilation. Vidard et al. (2007) found that when observations in the tropics were with-held, the system's background field changed enough to influence the quality control system's decisions. This led to data at higher latitudes being flagged as bad, ultimately degrading the model fields at higher latitudes. Several other instances of quality control decisions influencing OSE results in similar ways have been reported in the literature (e.g., Bouttier and Kelly 2006; Tremolet 2008). Subtleties like these can, in some cases, make OSEs difficult to interpret.

OSEs are usually conducted for a past period of time – for example, the last 3 years, or the time period when four satellite altimeters were operating. While this is very instructive, the GOOS is constantly changing (e.g., Figure 2). The number and distribution of Argo floats changes as new floats are deployed and old floats expire. New altimeter missions are launched and old missions end – and the sampling strategies of different altimeter missions are often different. This means that OSEs can become outdated. For example, using a seasonal forecast system, Vidard et al. (2007) and Balmaseda et al. (2007) perform a series of OSEs to evaluate the impact of Argo, XBT, and tropical moorings on forecast skill. Vidard et al. (2007) perform OSEs for the period 1993-2003 and Balmaseda perform OSEs for the period 2001-2006. So, for most of Vidard et al.'s OSEs, Argo coverage is sparse, while for most of Balmaseda et al.'s OSEs, Argo is substantial. As a result, Vidard et al. report only faint praise for the benefit of Argo, but note that it was probably too early to be sure. By contrast, Balmaseda et al. conclude that Argo is instrumental in iniatialising their forecast system – particularly for salinity.

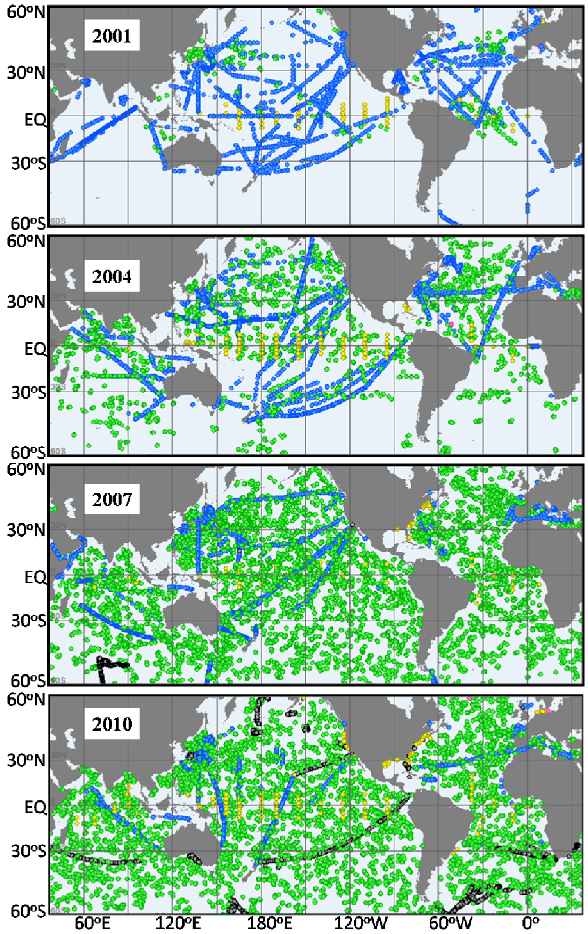

Figure 2: Observations during January of 2001, 2004, 2007, and 2010; green, blue, and yellow dots denote Argo floats, XBT/CTD profiles, and buoys respectively. Images sourced from www.coriolis.eu.org in February 2010.

Another limitation of OSEs is the significant computational and human resources required to undertake, analyse, and interpret them. Consider the study of Oke and Schiller (2007), for example. They conducted a series of 6-month model runs including an experiment with no assimilation, an experiment with all data assimilated, plus experiments with each observation type (Argo, SST, and altimeter) with-held. Additional experiments could include those with 1 altimeter, 2 altimeters, 3 altimeters, or 4 altimeters; experiments with different SST products assimilated; experiments with only a sub-set of Argo profiles, for example every other Argo profile. Such a series of OSEs equates to a significant amount of computation, and a large amount of data that requires processing, analysis, and interpretation. This is not always achievable, especially when a high resolution model is used.

Evaluation of OSEs is always a challenge. For any series of OSEs, the *best* experiment, by which all others are typically compared, is always the run that assimilates all observations. Evaluation of this run is therefore problematic, as there is usually no independent set of observations that can be used to evaluate this run.

An example of a series of OSEs, designed to evaluate the relative importance of altimetry, Argo, and SST for constraining an eddy-resolving ocean model, is described by Oke and Schiller (2007). Using a $1/10^{o}$ resolution ocean general circulation model and an ensemble optimal interpolation data assimilation system

(Oke et al. 2008), they systematically with-hold altimetry (denoted ALTIM), Argo, and SST from a reanalysis system for the period December 2005 to May 2006. The impact of with-holding each data type is illustrated in Figure 3, showing the residuals between reanalysed SLA and along-track SLA. The residual maps quantify the difference between observed and reanalysed SLA for each OSE. Reanalysed SLA is compared to along-track SLA from all available altimeters (Jason, Envisat, and GFO).

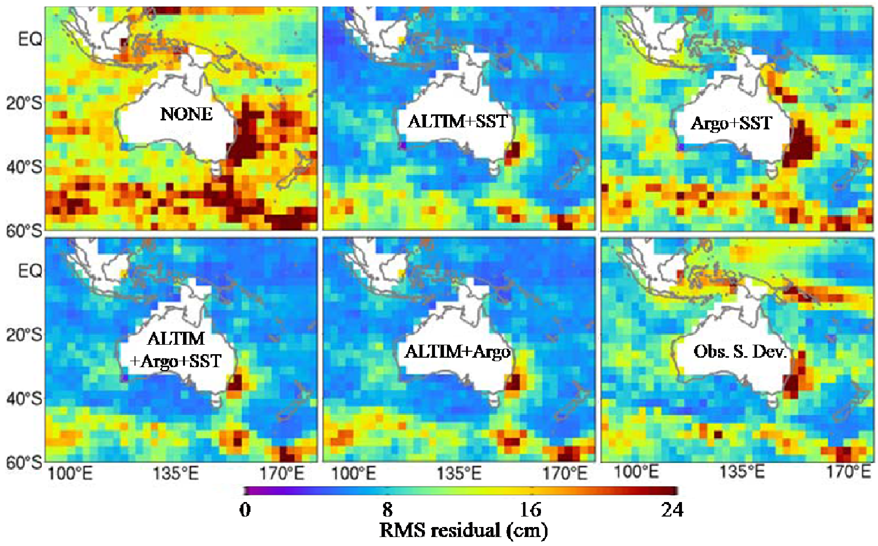

Figure 3: Root-mean-squared residual between modelled and observed sea-level anomaly for different OSEs. Adapted from Oke and Schiller (2007).

It results in Figure 3 indicate that when only Argo and SST are assimilated the SLA residuals are much smaller than the OSE that assimilated no observations, denoted NONE in Figure 3. This indicates that some of the information in altimetry is also represented by the SST and in situ T and S observations. This is expected, based on the well understood dynamical relationship between SLA and sub-surface T and S, but it also demonstrates the power of the multivariate EnOI scheme that is used by Oke and Schiller (2007). The SLA residuals are noticeably smaller when altimetry is assimilated, particularly in the regions of energetic mesoscale variability like the Tasman Sea, along the path of the Antarctic Circumpolar Current and off Western Australia, where the Leeuwin Current frequently sheds eddies (Figure 3). This suggests that while SST and Argo represent the broad-scale SLA features, they do not adequately resolve the details of the mesoscale.

Observing System Simulation Experiments - OSSEs

Another commonly used technique for evaluating the potential benefit of different observing systems is OSSEs. OSSEs often involve some sort of twin experiment, where *synthetic observations*, usually extracted from a model, are assimilated into an alternative model or gridded using an observation-based analysis system. OSSEs are commonly used to assess the impact of some hypothetical array of observations that may not exist yet. This means that these methods can be used to contribute to the design of future observing systems, quantifying their possible impacts and limitations.

OSSEs have been employed to support the design of oceanic observing systems since before the altimeter era. For example, the Berry and Marshall (1989), Holland and Malanotte-Rizzoli (1989), performed OSSEs to support the assessment of designs for the early altimeter missions. Similarly, OSSEs were conducted to support the design and assessment of the TAO array in the tropical Pacific Ocean (e.g., Miller 1990) and the PIRATA array in the tropical Atlantic Ocean (e.g., Hackert et al. 1998).

Several good examples of OSSEs were conducted during the planning of the tropical Indian Ocean mooring array (CLIVAR-GOOS Indian Ocean Panel and Co-authors 2006). These OSSEs were conducted by several different groups, using different models and different techniques. The results from these studies contributed to discussions during the planning of this mooring array. Vecchi and Harrison (2007) presented results from a series of OSSEs using a high resolution ocean model and an adjoint-based assimilation system to evaluate the ability of an integrated observing system, including Argo observations, XBT lines, and the proposed mooring array to monitor intraseasonal and interannual variability. Ballabrera-Poy et al. (2007) used a reduced-order Kalman filter to objectively determine an array for mapping sea surface height and SST. Oke and Schiller (2007) use an approach based on empirical orthogonal functions (EOFs) to assess the proposed mooring array's ability to monitor intraseasonal and interannual variability. Vecchi and Harrison (2007) concluded that in conjunction with the integrated observing system, the proposed mooring array should be capable of resolving intraseasonal and interannual variability. Both Ballabrera-Poy et al. (2007) and Oke and Schiller (2007) argued that the proposed array may oversample the region within a few degrees of the equator. These studies also suggested that key regions for monitoring seasonal- to-interannual variability are south of 8°S, at about 4°-5° from the equator and along the coast of Indonesia. These regions correspond to the locations of the maximum amplitude of seasonal Rossby waves (Masumoto and Meyers 1998; Schouten et al. 2002), equatorial Rossby waves, and strong Indian Ocean dipole events (Murtugudde et al. 2000), respectively.

An example of the above-mentioned OSSEs is presented in Figure 4, showing the standard deviation of the depth of the 20^{o}C isotherm (D20) from a model, along with the root-mean-squared error of D20 in two OSSEs. Each OSSE uses output from 18-years of a model run. The first 9-years are used to *train* the EOF-based analysis system that is described by Oke and Schiller (2007), and the last 9-years is used for cross-validation, and to evaluate how well different mooring arrays resolve variability of D20. For each OSSE, the last 9-years of the model run are sampled at mooring locations; those *observations* are perturbed with white noise according to their assumed errors; the *observations* are analysed, and the errors of the analysed D20 fields assessed. Figure 4 indicates that the proposed array resolves the variability of D20 very well near the equator, where the root-mean-squared errors are small, but poorly south of 10^{o}S, where the errors are relatively large. An alternative mooring array is also tested by Oke and Schiller (2007). The alternative array is generated objectively, by maximising the projection of observations onto an ensemble that is used for assimilation. For a detailed description of the method, the reader is referred to Oke and Schiller (2007). The alternative array, presented in Figure 4c, has fewer moorings close to the equator and in the northern Indian Ocean, and more moorings between 10-15^{o}S. Variability of D20 is still well resolved by the alternative array near the equator, but owing to the additional moorings to the south, the variability of D20 is better resolved there. The latitudes of high D20 variability to the south (10-15^{o}S) correspond to the maximum amplitude of seasonal Rossby waves (Masumoto and Meyers 1998; Schouten et al. 2002). The study by Oke and Schiller (2007) concluded by suggesting that additional moorings in those latitudes are worth considering.

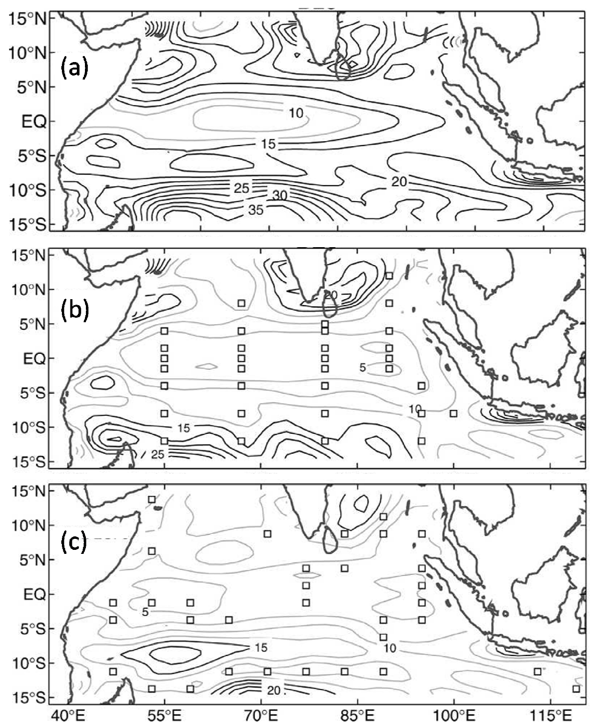

Figure 4: (a) Standard deviation of the depth of the 20^{o}C isotherm and the root-mean-squared error for (b) the proposed Indian Ocean Mooring array and (c) an optimized mooring array for a series of OSSEs; contour intervals are 2.5 m. Adapted from Oke and Schiller (2007).

OSSEs can be very instructive for assessing the potential impact of different observing systems. However, they have several limitations. It could be fair to say that OSSEs, in the form of twin experiments are *doomed to succeed* - particularly if the same model is used to produce the synthetic observations, as the model used for assimilation. In this case, the dynamics of the model and *observations* are perfectly compatible. As a result, some OSSEs using twin experiments report very low errors in assimilating model runs. In some cases, the errors are so low, and the results so optimistic, that the conclusions of such studies must be regarded with suspicion.

The relevance of any series of OSSEs ultimately depends on the assumptions made in configuring the OSSEs. In all cases, assumptions are made about the dynamics and the data assimilation methodology. It is implicitly assumed that the models capture the dynamics correctly and the observations are assimilated appropriately. Assumptions are made about the observation errors, and about model errors. In most cases, synthetic observations are corrupted by noise – and the noise is almost always assumed to be white in time and unbiased. It is also common to assume that there will be no data outages, and that data are all available at the time of assimilation. In the operational environment, this final assumption is rarely true.

Many OSSE studies employ methods that do not involve twin experiments. For example, Brassington and Divakaran (2009) analyse the theoretical impact of sea-surface salinity observations on an ensemble-based data assimilation system by examining various characteristics of the ensemble. Schiller et al. (2004) examine modelled fields to quantify the likely signal-to-noise ratios of different sampling strategies for the Argo program.

OSSEs can be a very instructive tool for evaluating the potential value of future observing systems. However, the assumptions made by OSSEs are often optimistic, and the results from OSSEs are therefore often optimistic – and should be regarded as indicative only, and perhaps qualitative in most cases.

Analysis Self-Sensitivity

In general, regardless of the method, a data assimilation system combines a background field (of either 2-, 3- or 4-dimensions) with a set of observations, yielding an analysis. Different assimilation methods do this in different ways. But for all methods, there exists a so-called analysis self-sensitivity. The analysis self-sensitivity quantifies the importance of each individual observation for a given analysis. Consider a couple of cases. Suppose we can change a given observation and the analysis does not change. In this case, we can say that the analysis is not sensitive to that observation, and conclude that it is unimportant. This may occur if the observation has a large error, or is in a region of dense observations – so it is redundant. Conversely, consider a case where a change to a given observation results in a significant change to the analysis. In this case, we can say that the analysis is sensitive to that observation, and conclude that it is important. This may occur if the observation is very accurate, or is in a data-sparse region. The sensitivity referred to above is called the analysis self-sensitivity.

In practice, self-sensitivities are diagnosed by the so-called influence matrix (Cardinali et al. 2004). The influence matrix is simply a subset of the Kalman gain, **K**. The Kalman gain is like a regression matrix, mapping each element of the background innovation (difference between a background field and observations) onto the full model state. The influence matrix is simply **HK**, where **H**, is an operator that interpolates from model-space to observation-space (often just linear interpolation). The matrix **HK** is square, with dimension *p* by *p*, where *p* is the number of observations assimilated. The diagonal elements of **HK** are the analysis self-sensitivities – they map the background innovation from the observation location to itself. Cardinali et al. (2004) and Chapnik et al. (2006) provide a practical recipe for diagnosing analysis self-sensitivities from any assimilation system – regardless of the assimilation method. Briefly,

1. Perform a standard analysis by assimilating observations **d**;
2. Perturb the assimilated observations ($\mathbf{d} \rightarrow \mathbf{d}^*$) according to their expected error (from the diagonal elements of **R**, the observation error covariance matrix);
3. Perform another analysis by assimilating the perturbed observations; and
4. Compute the self-sensitivities $\mathbf{HK_{ii}}$:

$$\mathbf{HK}_{ii} = (\mathbf{d}_i^* - \mathbf{d}_i)\,(\mathbf{Ha}_i^* - \mathbf{Ha}_i)\,/\,\mathbf{R_{ii}},$$

where **a** and $\mathbf{a}^*$ are analyses produced using unperturbed and perturbed observations respectively. The minimum calculation to estimate the self-sensitivities is a second analysis. However, this calculation is subject to sampling error, due to the random nature of perturbing the observations, so multiple perturbed analyses should be calculated in practice, to obtain robust estimates of the true self-sensitivities. The diagonals of the influence matrix can be analysed, or the partial trace of **HK** can be averaged for different regions, different variables, and so on.

With an estimate of the self-sensitivities at hand, it is common to diagnose the so-called degrees of freedom of signal (DFS) and the information content (IC) for different sub-sets of observations. The DFS provide an indication of how many truly independent observations are present in a given sub-set of observations. At most, the DFS is the same as the number of observations. In this case, the IC is 100% and there are no redundant observations. Conversely, if the DFS is much less than the number of observations, the IC of that set of observations is low. In this case, the IC may be small and there is significant redundancy in the observations.

An example of the IC and DFS for different observation types using the Bluelink reanalysis system (Oke et al. 2008) is given in Figure 5. 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 either not extracted by the Bluelink system in an optimal way, or is somewhat redundant – possibly well represented by the other assimilated observations. At this stage of development, the former explanation seems most likely. By producing these, and other, diagnostics from a number of GODAE systems, it is anticipated that the true value of all observations for GODAE systems can be routinely monitored and quantified. In turn, these evaluations could be fed back to the broader community for consideration.

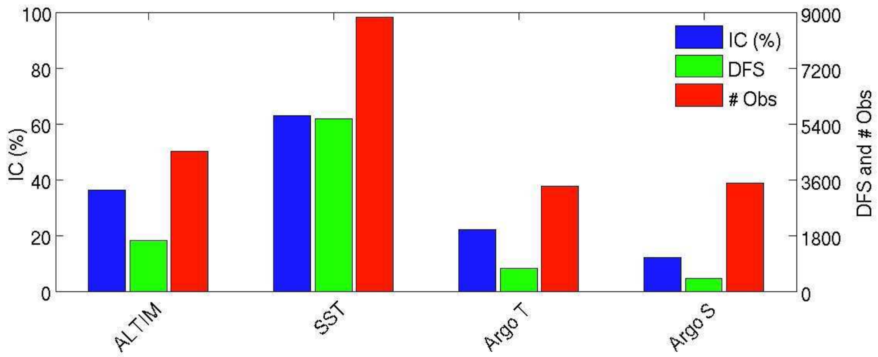

Figure 5: The Preliminary estimates of the Information Content (IC; %), degrees of freedom of signal (DFS) and the number of assimilated super-observations (# Obs) for the Bluelink reanalysis system in the region 90-180^{o}E, 60^{o}S-equator, computed for 1 January 2006. The scale for the IC is to the left and the scale for the DFS and # Obs is to the right.

In addition to providing a quantitative indication of the importance of each observation, and each observation type, for a given analysis, analysis self-sensitivities can be instructive for tuning assimilation and forecast systems. The goal of every assimilation system is to extract as much relevant information from every observation as possible. That is, to maximise the IC from the above-mentioned analysis. The type of diagnostic described here can contribute to this process.

Analysis self-sensitivity is a relatively inexpensive to perform and may be feasible for routine application to operational forecast systems. The latter point means that calculations could be performed on the modern-day GOOS. Limitations of analysis self-sensitivities however, include the fact that they are relevant only to analysis fields – not the forecast fields. Finally, self-sensitivities also depend on error estimates used by the assimilation or analysis system.

Ensemble-based methods

A variety of ensemble-based methods can be readily used for observing system design and assessment. These include the diagnosis of ensemble-based covariance fields, of which Figure 1 is an example, the objective ranking of the importance of observations with regard to their potential impact to minimise a system's analysis error variance, and diagnosis of bred vectors. A description and examples of these follow. Some good references for ensemble-based observing system design and assessment activities include Tracton and Kalnay (1993), Houtekamer and Derome (1995), Toth and Kalnay (1997), Bishop et al. (2001; 2003), and Wang and Bishop (2003).

An example of a series of ensemble-based correlation fields between sea-level at time t=0 days, and sea-level in the surrounding region 4-days earlier (t=-4 days) and 4-days later (t=+4 days) in the open ocean, south-west of New Caledonia is shown in Figure 6. The correlation fields provide insight into the underlying dynamics, the spatial length-scales, and the temporal length-scales of sea-level. For this example, a modified version of the 120-member stationary ensemble that is used by the Bluelink forecast and reanalysis system (Brassington et al. 2007; Oke et al. 2005; 2008) is used. It is evident that in this region a dominant dynamical process is the westward propagation of sea-level anomalies, probably characteristic of Rossby waves. The ensemble-based correlations indicate that the length-scales in this region are fairly short, with the influence of sea-level limited to within a few hundred kilometers of an observation. However, the time-scale seems to be quite long – the lagged correlations, for t=-4 and 4 days, are not very much less than the zero-lag correlations, for t=0 days (Figure 6). We therefore expect that an observation at some point in time is likely to be representative of the circulation for some time into the future and into the past. These factors may influence discussions on the appropriate spatial density and temporal sampling of observing systems in this region. Although the example presented in Figure 6 uses a stationary ensemble, and is therefore appropriate for the design and assessment of long-term monitoring programs, a time-evolving ensemble from an ensemble Kalman Filter system (e.g., Evensen 2003) that reflects the time- and state-dependent background field errors (so-called errors of the day; Corazza et al. 2003) could equally be used for adaptive sampling programs – where we might seek to identify good locations for imminent deployments of instruments, like gliders or profiling floats.

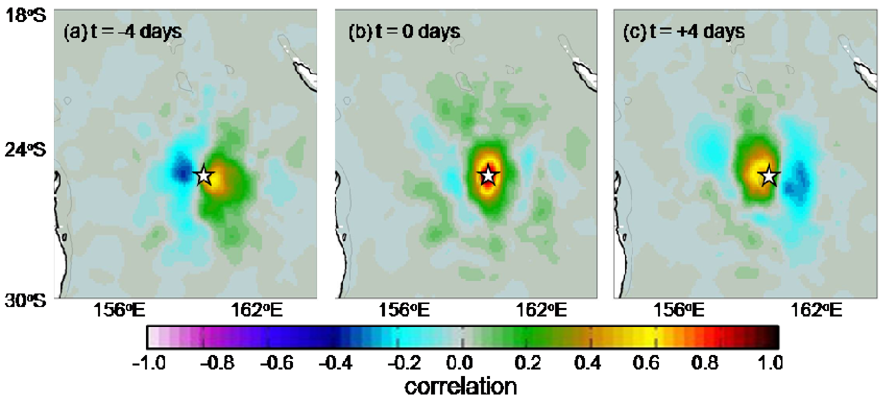

Figure 6: An example of four-dimensional ensemble-based correlation fields showing the spatio-temporal influence of a sea-level observation in the open ocean, south-west of New Caledonia. Each panel shows the ensemble-based correlations between sea-level at t = 0 days and sea-level in the surrounding region for time-lags of (a) -4 days, (b) 0 days, and (c) +4 days.

Ensemble-based methods for optimal array design are increasingly being used for NWP systems (e.g., Bishop et al. 2001). These methods are based on ensemble square root filter theory (e.g., Tippett et al. 2003; O'Kane and Frederiksen 2008b) and allow one to handle large systems in cases when explicit manipulation of the background error covariance matrix is not feasible. Most of the studies on the ensemble-based optimal array design consider the problem of adaptive sampling and targeted observation, aimed at improving the model's forecast at a given time (e.g., Bishop et al. 2001; Langland 2005; Khare and Anderson 2006).

The main steps in the ensemble-based objective design of an observation array are represented schematically in Figure 7. The first step is the construction of an initial ensemble that represents the system's background error covariance before any observations are assimilated. Such an ensemble might be associated with some variant of the ensemble Kalman Filter (Evensen 2003). Given an ensemble that implicitly represents the system's background error covariance, and an array of observations of known error variance, ensemble square root theory provides an efficient framework for updating, or transforming, the ensemble so that its updated error variance matches the theoretical analysis error variance after those observations are assimilated (Bishop et al. 2001). There are several ways of implementing this transformation (see Tippett et al. 2003), all of which are equivalent, but the most computationally efficient transformation is that of the ensemble transform Kalman filter (ETKF; Bishop et al. 2001) and specifically the serial implementation of the ETKF. So, the second step is to update the ensemble to represent the system's error covariance after assimilation of all available observations (Figure 7). The third step is to identify the *next best* targeted observation. That is, the observation that transforms the ensemble to yield the ensemble with the smallest analysis error variance. This targeted observation is identified by explicitly transforming the ensemble for all possible observations and identifying the observation that minimises the ensemble's analysis error variance. This is a *brute force* calculation – however, the update from a single observation is inexpensive, so this approach is generally feasible, even for systems with a large state dimension. Once the latest targeted observation has been identified, the ensemble is updated and the process of identifying the next best targeted observation is repeated, until the number of targeted observations has been reached.

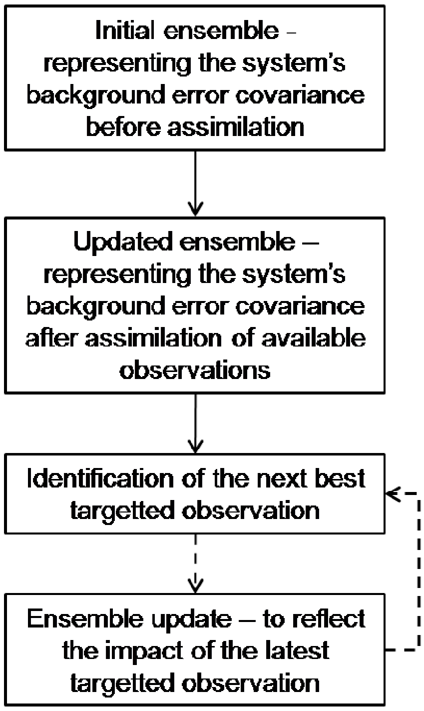

Figure 7: Schematic diagram depicting the serial calculation of an optimal observation array. The dashed arrows represent the serial identification of targeted observations and the ensemble updates that reduce the ensemble's variance given those targeted observation. Adapted from Sakov and Oke 2008.

The most important step in the ensemble-based approach described above is the determination of what the targeted observations are intended to minimise. In practice, the ensemble includes several different variables (e.g., temperature, salinity, velocity etc.). The identification of the next best targetted observation can be performed so that it minimises a specific aspect of the analysis error. For example, it might minimise the analysis errors of temperature in a specific target region, or the analysis errors of mixed layer depth, or the volume transport through a strait of passage. This criterion may have a significant impact on the objectively designed observation array (e.g., Sakov and Oke 2008). Careful determination of what is to be minimised is important. For this to be achieved, it is important to be very clear about the purpose, or motivation, of the observation array.

An example of an ensemble-based objective observing system design, from Sakov and Oke (2008), is presented in Figure 8. This example addresses the design of the tropical Indian Ocean mooring array. It is assumed that the purpose of this array is to minimise the analysis error variance of Intraseasonal Mixed Layer Depth (IMLD). Figure 8 shows the error variance of IMLD before and after assimilation, for two different models and for three different mooring arrays, and assumes that no other observations are available (i.e., no Argo, XBT, altimeter data etc.). It is assumed that observations from the mooring array are to be assimilated into a model using an ensemble-based data assimilation system using a stationary ensemble. Two different ensembles are considered, each generated by different model configurations (ACOM2 and ACOM3), with different forcing, and integrated for different periods. Three different options for the mooring array are considered: the proposed mooring array (denoted CG-IOP array), and an optimised array for each model, denoted ACOM2-array and ACOM3-array. In this case, the initial ensemble variance for IMLD is shown, along with the final ensemble variance for IMLD given the different mooring arrays (Figure 8). Sakov and Oke (2008) use different models here in pursuit of more robust results.

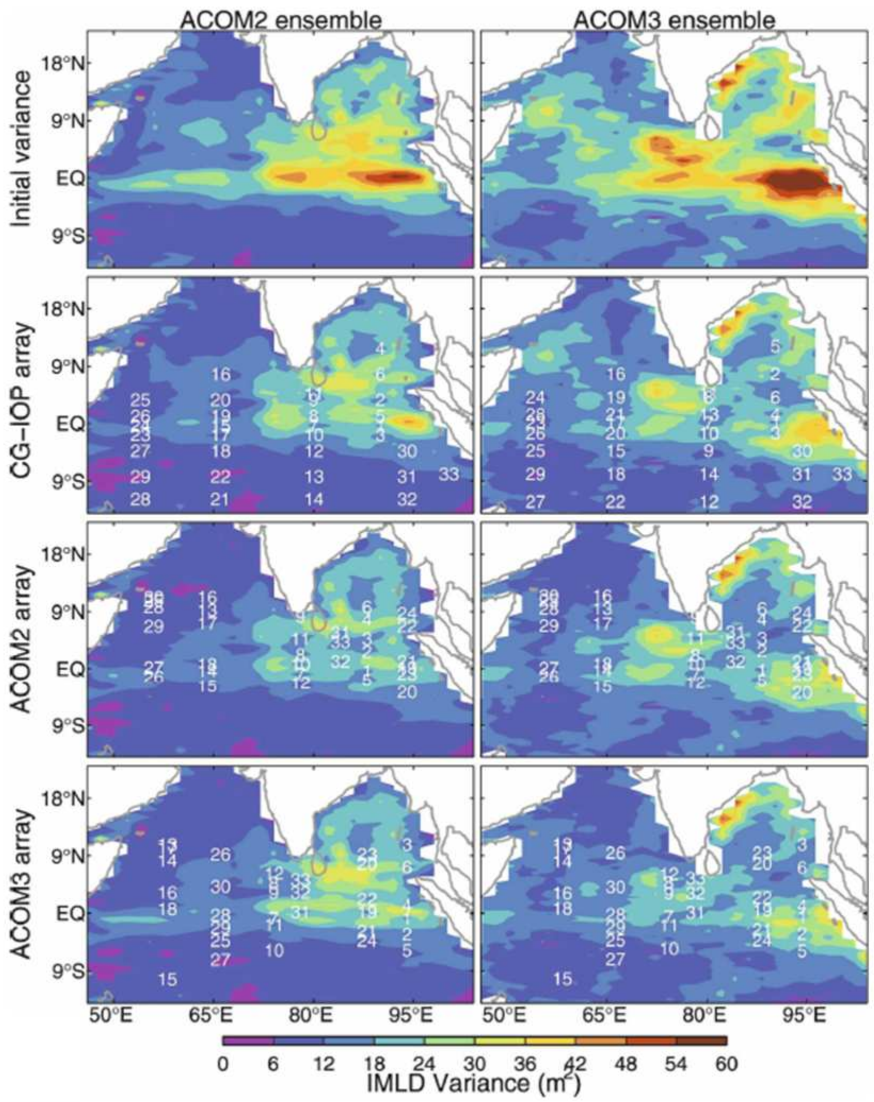

Figure 8: The variance of the IMLD (top row) in ACOM2 (left) and ACOM3 (right), and the theoretical analysis error variance for each model using the CG-IOP-array (2nd row), and the arrays derived using ensembles from ACOM2 (3rd row) and ACOM3 (4th row), as labelled to the left of each row. The numbers 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).

The numbers overlying the error variance maps in Figure 8 refer to an objective ranking of each observation location – the order in which they were identified by the method depicted in Figure 7. In each case, the mooring array is constrained to a limited number of mooring lines at distinct longitudes – to simplify routine maintenance of the array. Using the ETKF framework, the *best* mooring line is identified, and then the *best* observation for that mooring line is derived. So the mooring line with numbers 1-6 is the *best* mooring line. For each array considered and for both models, the best mooring line is located in the eastern Indian Ocean, between 90-95^{o}E, and the mooring line south of India is also very important, ranked 7-12 (or 7-14) for each scenario considered. These results appear to be robust, and can aid the decision-makers when mooring design and priorities are being made – for example, which mooring line should be deployed first?

Breeding is an ensemble technique that seeks to quantify the structures of the fastest-growing dynamical modes of a model. Bred vectors are perturbations to the model state that grow rapidly in time. Bred vectors are particularly useful for adaptive sampling, where the errors of the day are used to identify where an instability is most likely to originate. More observations in a region of instability might better constrain a deterministic forecast, resulting in better forecast skill.

Breeding was first explored by Toth and Kalnay (1997) for an NWP ensemble prediction system. In practice, bred vectors are generated by first initialising a model with an ensemble of perturbations. Initially, the perturbations are typically simply small-amplitude white noise. The ensemble is integrated for a fixed period of time. The perturbations are periodically rescaled using a global (or regional) scale factor so that they approximate fast-growing errors within an assimilation scheme. The choice of scale factor is important. One of the purposes of breeding is to identify fast-growing instabilities (O'Kane and Frederiksen 2008c). In some regions, these instabilities will be best represented by sea-level anomalies; in other regions in might be sub-surface temperature, or density. This should be tuned for different regions. However, some atmospheric applications have demonstrated that the choice of rescaling doesn't significantly influence the bred vectors (e.g., Corazza et al. 2003). This is in contrast to singular vectors (see below), which are very sensitive to the choice of norm (e.g., Palmer et al. 1998; Snyder et al. 1998).

In practice, the ensemble perturbations (bred vectors) usually become well-organised, coherent structures that can be interpreted and understood (e.g., instabilities associated with an eddy). This approach readily allows ensembles to be initialised about the analysis from data assimilation that contain, by construction, information about the errors of the day. Thus the bred vectors tend to project strongly onto regions where forecast errors are large. The process of breeding is represented schematically in Figure 9.

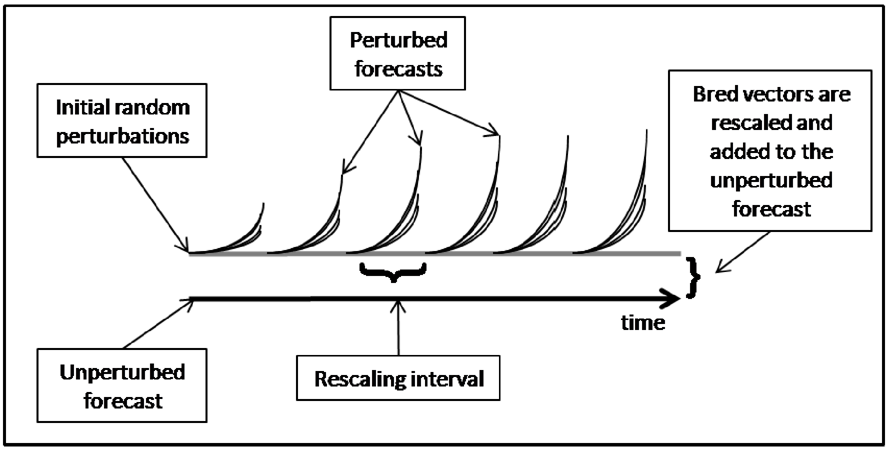

Figure 9: Schematic diagram depicting the generation of bred vectors. An ensemble is initially perturbed with uncorrelated noise. The rescaling parameter must be chosen carefully (e.g., temperature at 250 m depth in key region). After each rescaling interval the ensemble perturbations are rescaled to the same magnitude as the initial perturbations – but bred vectors develop spatially coherent, well-organised structures. Each bred vector is the difference between a perturbed forecast and the unperturbed forecast.

For an atmospheric example, Houtekamer and Derome (1995) showed that bred vectors produce similar results to singular vectors (described below), but they are much easier to implement (Wei and Frederiksen 2004). Because of its simplicity, breeding is a very versatile approach. Bred vectors have recently been explored by many operational global weather predictions systems (e.g., O'Kane et al. 2008) using an implementation that is based on the ETKF (e.g., Wang and Bishop 2003; Wei et al., 2006). The ETKF is a generalisation of breeding, but it is more complex and more computationally expensive. The main difference is that the ETKF orthogonalises the bred vectors and seeks to maximise the ensemble spread.

An example of breeding applied to a regional ocean model of the Tasman Sea is presented in Figure 10. For this example, a 4-member ensemble is used, and bred vectors are optimized (rescaled to) amplify temperature anomalies at 250 m depth. The forecast errors for sea-level, computed by comparing with a verifying analysis, are shown in Figure 10a-d along with the 4-member ensemble averaged bred vector overlaid. The individual bred vectors are also contoured in Figure 10e-h. For the period shown here, the forecast error for sea-level is quite large at several locations. The bred vectors are independent of the forecast error; however they project strongly onto the regions where the forecast error is large and spatially coherent. This indicates that the bred vectors are reliably identifying regions of growing instabilities. For the case displayed in Figure 10, the forecast does not pick up the developing instabilities (see 11 March) that, in this case, correspond to a developing cold-core eddy. With regard to adaptive sampling, if this 4-member breeding system is run in parallel with the operational forecast system, the regions of strong growth in the bred vectors might be good candidates for the deployment of additional observations – perhaps in the form of gliders or profiling floats. In this case, the improved initialisation of the forecast in those regions might have better constrained the forecast and improved the forecast skill for this event.

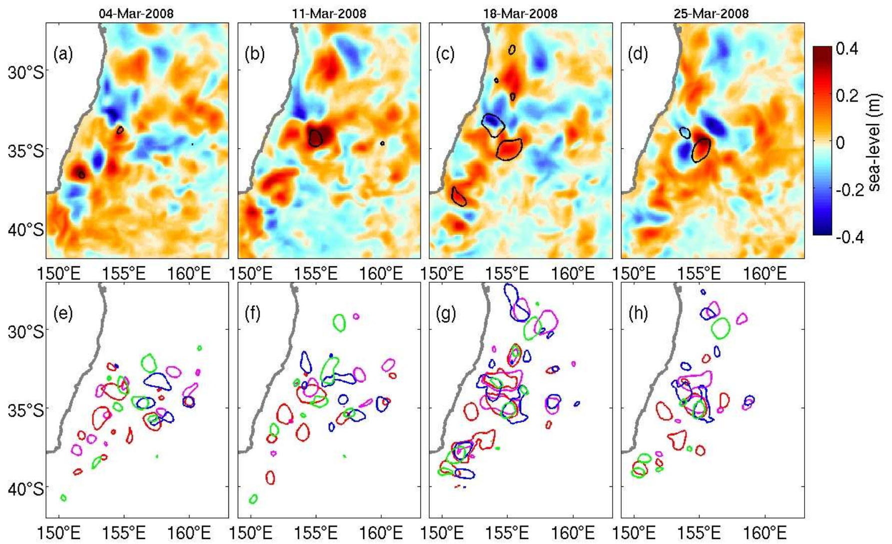

Figure 10: Examples of (a-d) forecast error for sea-level (colour) and the ensemble averaged bred vector (contours) in the Tasman Sea; and (e-h) four bred vectors overlaid. Each bred vector is a different colour.

Adjoint-based methods

A variety of adjoint-based methods can be readily used for observing system design and assessment. These include diagnosis of representers, adjoint sensitivities, and singular vectors. A description and examples of these follow. Some good references for adjoint-based observing system design and assessment activities include Moore and Farrell (1993), Rabier et al. (1996), Gelaro et al. (1998), Palmer et al. (1998), Baker and Daley (2000), Langland and Baker (2004), and Moore et al. (2009).

Representers are analogous to the ensemble-based covariance fields displayed in Figure 1 and Figure 6. Representers quantify the temporal and spatial *footprints* *of influence* of an observation. Using the system's tangent linear model to trace the influence of an observation into the future, and its adjoint to trace its influence into the past, an adjoint-based data assimilation system readily approximates the covariance between a given observation location and type (e.g., sea-level at a fixed location) and all other variables at all model grid locations for all time. Representers can help build intuition about how different observation types and locations influence a data assimilating model. An example of the components of a representer, derived from the Advanced Variational Regional Ocean Representer Analyzer (AVRORA) system (Kurapov et al. 2009), for the coastal ocean is presented in Figure 11. The background field for these calculations corresponds to an idealised two-dimensional wind-driven upwelling scenario (Figure 11a) with characteristics of the upwelling circulation off Oregon, USA. Details of the model configuration and assimilation system are described by Kurapov et al. (2009). They investigate the structure of representers to better understand the potential impact of assimilating observed sea-level anomalies from altimeters into a coastal ocean model.

The representer components shown in Figure 11 quantify the covariance between a hypothetical sea-level observation that is 50 km offshore, and the rest of the model state. The components shown in Figure 11 are for the time of the observation. The full representer includes time, where the influence of the observation extends over both time and space.

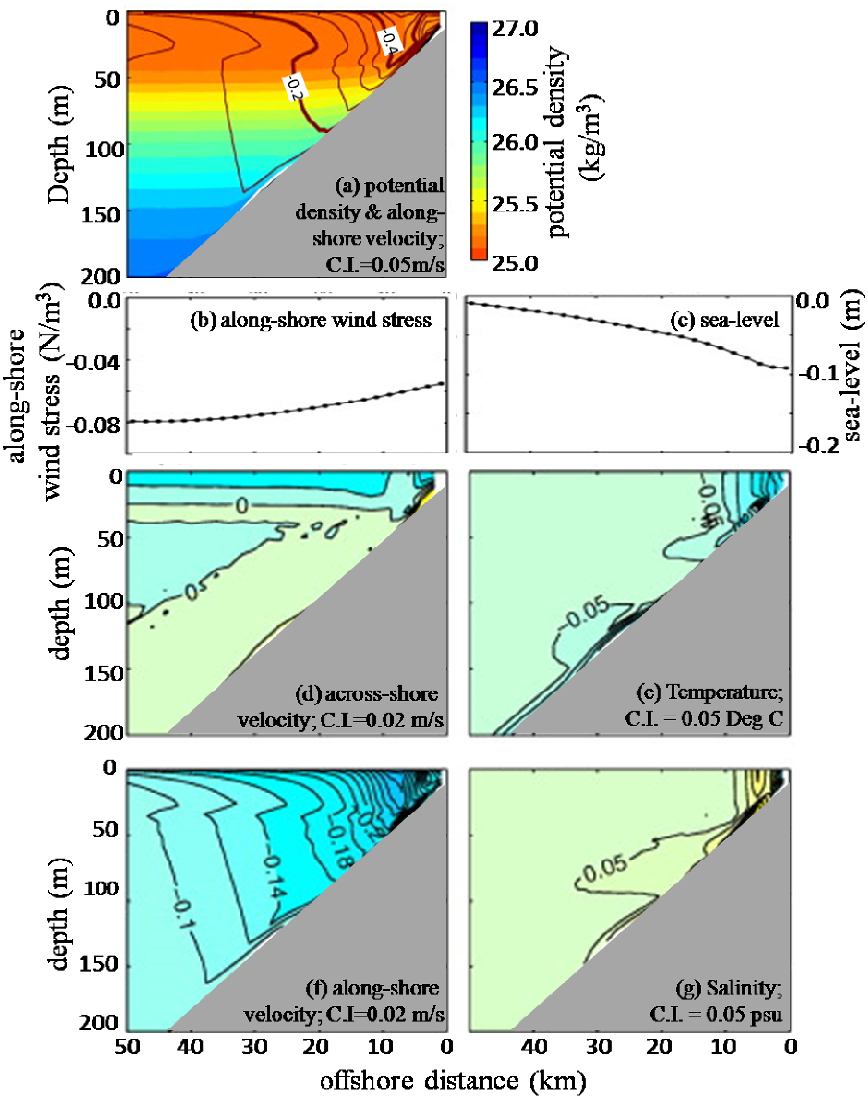

Figure 11: The components of a representer in a cross-shore section for an idealised two-dimensional wind-driven upwelling scenario (panel (a) shows the background field), showing the covariance at the time of the observation (zero time-lag) between sea-level 20 km from shore and (b) along-shore wind-stress, (c) sea-level, (d) across-shore velocity, (e) temperature, (f) along-shore velocity, and (g) salinity. Contour intervals are provided in the titles for each plot. The contour intervals (C.I.) for panels (a, d-g) are marked. Adapted from Kurapov et al. 2009.

The fields in Figure 11 show how the assimilation system updates the model state when the observed offshore sea-level is lower than the modelled background estimate. The changes introduced by the assimilation are consistent with a strengthening wind-driven upwelling, with stronger upwelling favourable along-shore wind stress, lower sea-level over the shelf, offshore flow near the surface and on-shore flow through a bottom boundary layer, an accelerated baroclinic coastal jet, and a temperature (salinity) decrease (increase). The representer fields presented in Figure 11 indicates that offshore sea-level observations from altimetry are suitable for assimilation into coastal ocean models, and are likely to impose a significant constraint on the circulation over the continental shelf.

Adjoint, or observation, sensitivities seek to quantify the sensitivity of a forecast to assimilated observations (Langland and Baker 2004). Specifically, adjoint sensitivity determines the sensitivity of the cost function J, with respect to each observation y: that is, dJ/dy. In practice, Langland and Baker (2004) provide a practical recipe for computing adjoint sensitivities as follows:

1. Define the error norm of interest (e.g., position of an eddy, or the variability of a given variable in a region of interest);
2. Perform a forecast from say, t=0 to t=7, where t denotes time;
3. Compute a verifying analysis for $t=7$ (not in real-time);
4. Compute the difference between a forecast (valid at t=7) and a verifying analysis (also valid at t=7). This difference is an estimate of the forecast error.

5. Initialise the adjoint model with the forecast error and integrate backwards (from T=7 to t=0), yielding a new initial condition (valid at t=0); and
6. Calculate the sensitivity of the forecast to each observation, or a subset thereof.

Like all variational data assimilation tools, a model's tangent linear version and the adjoint of its tangent linear model are required to perform adjoint sensitivities. However, the adjoint 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. Like analysis self-sensitivities, described above, adjoint sensitivities can help identify low-influence and high-influence observations; and can be partitioned for any data subset: instrument type, observed variable, geographic region, vertical level, or individual reporting platform; thereby making the diagnostic directly relevant to GOOS data providers. Importantly, both analysis and adjoint sensitivities do not necessarily quantify the value of the observations – rather they quantify how much of the observations are used by an assimilation and forecast system given the assumed error estimates therein.

Like bred vectors, singular vectors are the fastest growing perturbations for a specific region at a specific time, and are most suited for adaptive sampling (e.g., Baker and Daley 2000). Unlike bred vectors, singular vectors are assumed to grow linearly in time. Singular vectors are perturbations with the greatest linear growth over a specified time interval, for a given norm, and defined over a specified target area. Singular vectors are only valid for time intervals for which the growth of a perturbation is linear. For the atmosphere, this is likely to be limited to a few days, and for the ocean it is possibly a week or two, depending on the underlying dynamics. To determine the growth of a perturbation over time a tangent-linear version of the full non-linear forecast model is required, along with the adjoint of the tangent linear model.

Before one can compute the fastest growing perturbations an appropriate choice of norm must be made for each. Ideally the initial norm is related to the spatial distribution of expected errors in the analysis while the final-time norm should reflect the forecast errors of interest. In practice, in NWP the total energy is often used for both initial and final time norms (e.g., at ECMWF). In practice mixed evolved and initial singular vectors are used in ensemble prediction allowing the growth rates of the perturbations to be tuned for a given application.

The notion of a target area is important for the computation of singular vectors. Singular vectors are the initial perturbations that result in the fastest growing perturbations in a target region. For example, Fujii et al. (2008) seek to predict the development of the Kuroshio meander with a lead time of 60 days. The target area is the region in which the Kuroshio typically meanders. Singular vectors are the perturbations, either within or outside of the target area, that result in large perturbations in the target area 60-days after initialisation. In NWP, the target area might be a major city and the time interval might be 10 days. The singular vectors are the initial perturbations that lead to large changes over that major city 10 days in the future.

Different choices of time interval, norm, or target area lead to different sets of singular vectors (e.g., Palmer et al. 1998; Snyder et al. 1998). This is in contrast to bred vectors that are relatively insensitive to the choice of rescaling (e.g., Corazza et al. 2003).

An example of an adjoint-based method used to calculate forecast sensitivity is described by Fujii et al. (2008). They use the Multivariate Ocean Variational Estimation system to investigate the types of perturbations that influence the large meanders in the Kuroshio Current. Specifically, they show that the leading singular vector represents a growing perturbation that leads to further development of the large meander. Figure 4a shows the perturbation to vertical velocity and pressure at 820-m depth at initial time. The anticyclonic anomaly positioned at 133°E, 31°N causes cold advection across the Kuroshio Current and downwelling to the north. This results in the development of an anticyclonic circulation in the deep layers, and induces baroclinic instability. The corresponding anomalies to sea surface height (SSH) that coincide with these developments are summarized in Figure 4b-d, showing the development of a large meander about two 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 133°E, 31°N and particularly at depths of 1000 to 1500 m. Thus, additional observations in that region are likely to benefit the forecast of the variability of the Kuroshio Current.

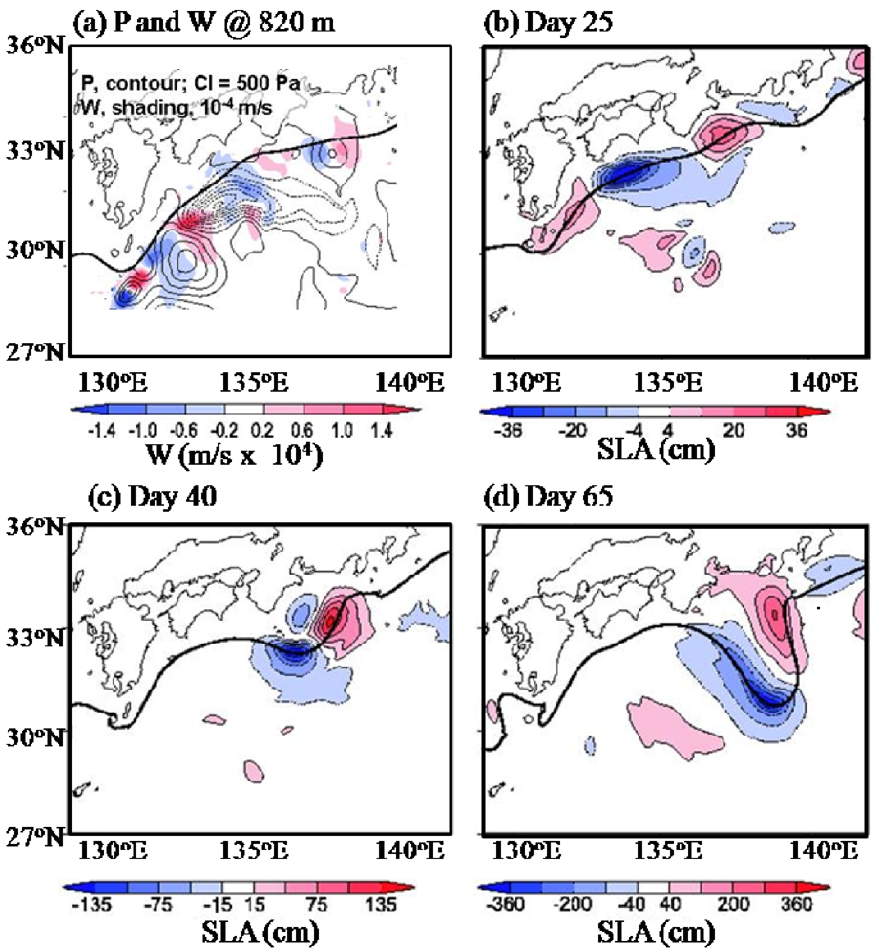

Figure 12: (a) Perturbation fields for pressure (contour; dotted lines are negative) and vertical velocity (shading; positive is downward) at 820-m depth. (b-d) SLA (scales are different for each panel) that result from the perturbations represented in panel (a) at Day 0. Thick lines show the Kuroshio Current axis in the background state. Adapted from Fujii et al. (2008a).

Summary

The use of models and data assimilation tools to aid the design and assessment of ocean observing systems is increasing. The most commonly used technique for evaluating the relative importance of existing observations is OSEs and OSSEs. OSEs are particularly useful for evaluating the relative importance of existing observations. But they are expensive to perform and analyse, and are sometimes difficult to evaluate and interpret. Despite this, as probably the simplest method for evaluating observing systems, OSEs are commonly used. OSSEs are most useful for examining the potential benefits of future observational platforms, and for contrasting the relative merits of different observational strategies. Like OSEs, OSSEs are easily implemented. However, OSSEs tend to return overly optimistic results, owing to the implicit dynamical consistency between the model-generated *observations* that are assimilated, and the models into which those observations are assimilated. Also, OSSEs are always limited by the realism of the models that are used.

Like OSEs and OSSEs, analysis self-sensitivities can be computed from an assimilation system regardless of the assimilation methods being used. Analysis self-sensitivities quantify the relative importance of every assimilated observation for a given implementation. Unlike OSEs, analysis self-sensitivities are relatively inexpensive to compute and analyse, and could feasibly be implemented routinely by operational centers. In this case, analysis sensitivities could provide an up-to-date, routine evaluation of the current observing system. Such analyses could be very beneficial to the observational community, by identifying existing and developing *gaps* in the global ocean observing system.

A range of ensemble-based techniques are available for observing system design and assessment. These include objective, ensemble-based array design (e.g., Sakov and Oke 2008), breeding (e.g., Toth and Kalnay 1997), and variants of breeding, like the ETKF (Bishop et al. 2001). Ensemble-based methods generally require an ensemble-based data assimilation system, such as ensemble optimal interpolation (e.g., Oke et al. 2008) or the ensemble Kalman Filter (Evensen 2003), for their application. Ensemble-based techniques are generally easily implemented, but often require significant computational resources and are subject to sampling error.

Various adjoint-based methods are also suitable for observing system design and assessment. These include analysis of representers (e.g., Kurapov et al. 2009), adjoint sensitivities (e.g., Langland and Baker 2004) and singular vectors (e.g., Fujii et al. 2008). The application of adjoint-based techniques generally requires a system's tangent linear model and its adjoint to be available.

Bred vectors and singular vectors are somewhat analogous. Both methods diagnose the system's fastest growing modes, or instabilities. With respect to adaptive sampling, regions where these modes project strongly might be good places to deploy additional observations. Assimilation of those additional observations may improve the initialisation of a forecast, thereby improving its forecast of the developing instability. Although bred vectors and singular vectors are very similar, in practice, breeding is much more easily implemented. Also the details of bred vectors are relatively insensitive to the details of the rescaling parameter, or norm, used in the breeding process, but bred vectors *are* sensitive to the rescaling interval. By contrast, singular vectors do tend to be sensitive to the choice of norms used.

The field of observing system design and assessment has seen many advances in techniques over the past decade. Together with the maturing nature of ocean forecasting, this has seen an increase in the use of models and data assimilation tools to aid the design and assessment of observing systems. The relevance of most methods depends on the realism of the models used. One way to combat this is to employ multiple methods and multiple models. Under the auspices of GODAE OceanView, it is hoped that this can be achieved through real international cooperation.

Acknowledgments

Financial support for this research is provided by CSIRO, the Bureau of Meteorology, and the Royal Australian Navy as part of the Bluelink project, and the US Office of Naval Research (Grant No. N00014-07-1-0422). Satellite altimetry is provided by NASA, NOAA, ESA and CNES. Drifter data are provided by NOAA-AOML and SST observations are provided by NASA, NOAA and Remote Sensing Systems. Argo data are provided by the Coriolis and USGODAE data centres.

References

Argo Science Team: On the design and implementation of Argo: An initial plan for a global array of profiling floats. International CLIVAR Project Office Rep. 21, GODAE Rep. 5, GODAE Project Office, Melbourne, Australia, 32 pp. (1998)

Baker, N.L., Daley, R.: Observation and background adjoint sensitivity in the adaptive observation targeting problem. Q. J. R. Meteorol. Soc. **126**, 1431–1454 (2000)

Ballabrera-Poy, J., Hackert, E., Murtugudde, R., Busalacchi, A.J.: An observing system simulation experiment for an optimal moored instrument array in the tropical Indian Ocean. J. Climate, **20,** 3284–3299 (2007)

Balmaseda, M.A., Anderson, D., Vidard, A.: Impact of Argo on analyses of the global ocean. Geophys. Res. Lett., **34** (2007) doi:10.1029/2007GL030452

Barth, N.H.: Oceanographic experiment design II: genetic algorithms, J. Atmos. Oceanic Tech. 9 (1992), pp. 434–443.

Berry, P., Marshall, J.: Ocean modelling studies in support of altimetry. Dyn. Atmos. Oceans, **13**, 269-300 (1989)

Bishop, C.H., Etherton, B.J., Majumdar, S.J.: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. Mon. Wea. Rev., **129,** 420–436 (2001)

Bishop, C.H., Reynolds, C.A., Tippett, M.K.: Optimization of the fixed global observing network in a simple model. J. Atmos. Sci., **60**, 1471-1489 (2003)

Bouttier, F., Kelly, G.: Observing-system experiments in the ECMWF 4D-Var data assimilation system. Q. J. Roy. Met. Soc., **127**, 1469 – 1488 (2006)

Brassington, G. B., Pugh, T., Spillman, C., Schulz, E., Beggs, H., Schiller, A., Oke, P.R.: BLUElink> Development of operational oceanography and servicing in Australia, J. Res. Practice Inf. Tech., **39**, 151-164 (2007)

Brassington, G.B., Divakaran, P.: The theoretical impact of remotely sensed sea surface salinity observations in a multi-variate assimilation system. Ocean Mod., **27**, 70-81 (2009)

Cardinali, C., Pezzulli, S., Andersson, E.: Influence-matrix diagnostic of a data assimilation system, Q. J. Roy. Met. Soc., **130**, 2767-2786 (2004)

Chambers, D.P., Tapley, D.B., Stewart, R.H.: Anomalous warming in the Indian Ocean coincident with El Niño. J. Geophys. Res., **104,** 3035–3047 (1999)

Chapnik, B., Desroziers, G., Rabier, F., Talagrand, O.: Diagnosis and tuning of observational error statistics in a quasi operational data assimilation setting. Q. J. Roy. Meteorol. Soc., **132**, 543–565 (2006)

CLIVAR–GOOS Indian Ocean Panel and Co-authors: Understanding the role of the Indian Ocean in the climate system - Implementation plan for sustained observations. WCRP Informal Rep. 5/2006, ICOP Publ. Series 100, GOOS Rep. 152, 76 pp (2006)

Corazza, M., Kalnay, E., Patil, D., Yang, S.-C., Morss, R., Cai, M., Szunyogh, I., Hunt, B., Yorke, J.: Use of the breeding technique to estimate the structure of the analysis errors of the day. Nonlin. Proc. Geophys. **10**, 233–243 (2003)

Evensen, G.: The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean Dyn., **53,** 343– 367 (2003)

Feng, M., Meyers, G.A., Wijffels, S.E.: Interannual upper ocean variability in the tropical Indian Ocean. Geophys. Res. Lett., **28,** 4151–4154 (2001)

Frederiksen, J.S., O'Kane, T.J.: Entropy, closures and subgrid modelling. Entropy, 10, 635-683 (2008)

Fujii, Y., Tsujino, H., Usui, N., Nakano, H., Kamachi, M.: Application of singular vector analysis to the Kuroshio large meander. J. Geophy. Res., **113** (2008). doi: 10.1029/2007JC004476.

Gallagher, K., Sambridge, M., Drijkoningen, G.: Genetic algorithms: an evolution from Monte-Carlo methods for strongly non-linear geophysical optimization problems, Geophys. Res. Lett., **18** 2177-2180 (1991)

Gelaro, R., Buizza, R., Palmer, T.N., Klinker, E.: Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors, J. Atmos. Sci., **55**, 1012–1037 (1998)

Gelaro, R., Langland, R.H., Rohaly, G.D., Rosmond, T.E.: As assessment of the singular-vector approach to targeted observing using the FASTEX dataset, Q. J. R. Meteorol. Soc., **125**, 3299– 3327 (1999)

Guinehut, S., Le Traon, P.-Y., Larnicol, G., S. Phillips, S.: Combining Argo and remote-sensing data to estimate the ocean three-dimensional temperature fields: A first approach based on simulated observations. J. Mar. Sys., **46**, 85–98. (2004)

Hackert, E.C., Miller, R.N., Busalacchi, A.J.: An optimized design for a moored instrument array in the tropical Atlantic Ocean. J. Geophys. Res., **103,** 7491–7509 (1998)

Heimbach, P., and Co-authors: Observational requirements for global-scale ocean climate analysis: Lessons from ocean state estimation. In Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21-25 September 2009, Hall, J., Harrison, D. E. & Stammar, D., Eds., ESA Publication WPP-306, submitted (2010)

Hernandez, F., Le Traon, P.-Y., Barth, N.: Optimizing a drifter cast strategy with a genetic algorithm, J. Atmos. Oceanic Tech., **12** 330-345 (1995)

Holland, W.R., Malanotte-Rizzoli, P.: Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies. J. Phys. Oceanog., **19**, 1507-1534 (1989)

Houtekamer, P., Derome, J.: Methods for ensemble prediction. Mon. Wea. Rev, **123**, 2181-2196 (1995)

Khare, S.P., Anderson, J.L.: An examination of ensemble filters based adaptive observation methodologies. Tellus, **58A,** 179–195 (2006)

Kuo, T.H., Zou, X., Huang, W.: The impact of global positioning system data on the prediction of an extratropical cyclone: An observing system simulation experiment. Dyn. Atmos. Oceans, **27,** 439–470 (1998)

Kurapov, A.L., Egbert, G.D., Allen, J.S., Miller, R.N.: Representer-based analyses in the coastal upwelling system. Dyn. Atmos.Oceans, 48, 198-218 (2009)

Langland, R.H., Baker, N.L.: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. Tellus, **56A,** 189–201 (2004)

Langland, R.H.: Issues in targeted observations. Quart. J. Roy. Meteor. Soc., **131,** 3409–3425 (2005)

Masumoto, Y., Meyers, G.A.: Forced Rossby waves in the southern tropical Indian Ocean. J. Geophys. Res., **103,** 27 589–27 602 (1998)

McPhaden, M.J., and Co-authors: The Tropical Ocean Global Atmosphere (TOGA) observing system: a decade of progress. J. Geophys. Res., **103**, 14169–14240 (1998)

Miller, R.N.: Tropical data assimilation experiments with simulated data: The impact of the Tropical Ocean, Global Atmosphere Thermal Array for the Ocean. J. Geophys. Res., **95**, 11461-11482 (1990)

Moore, A.M., Farrell F.:, Rapid perturbation growth on spatially and temporally varying oceanic flows determined using an adjoint method: Application to the Gulf Stream, J. Phys. Oceanogr., **23**, 1682– 1702 (1993)

Moore, A.M., Arango, H.G., Di Lorenzo, E., Miller, A.J., Cornuelle, B.D.: An adjoint sensitivity analysis of the southern California current circulation and ecosystem. J. Phys. Oceanogr., **39**, 702-720 (2009)

Murtugudde, R., McCreary, J.P., Busalacchi, A.J.: Oceanic processes associated with anomalous events in the Indian Ocean with relevance to 1997–1998. J. Geophys. Res., **105,** 3295–3306 (2000)

O'Kane, T.J., Frederiksen, J.S.: Statistical dynamical subgrid-scale parameterizations forgeophysical flows. Phys. Scr., T132, 014033, doi: 10.1088/0031-8949/2008/T132/014033 (2008a)

O'Kane, T.J.,, Frederiksen, J.S.: Comparison of statistical dynamical, square root and ensemble Kalman filters. Entropy, 10, 684-721 (2008b)

O'Kane, T.J., Frederiksen, J.S.: A comparison of statistical dynamical and ensemble prediction methods during blocking J. Atmos. Sci., 65, 426-446 (2008c)

O'Kane, T.J., Naughton, M., Xiao, Y.: AGREPS: The Australian Global and Regional Ensemble Prediction System. ANZIAM J., 50, C308-C321 (2008)

Oke, P.R., Schiller, A., Griffin, D.A., Brassington, G.B.: Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. Quart. J. Roy. Meteor. Soc., **131,** 3301–3311 (2005)

Oke, P.R., Schiller, A.: Impact of Argo, SST and altimeter data on an eddy-resolving ocean reanalysis. Geophys. Res. Lett., **34** (2007). doi: 10.1029/2007GL031549 (2007)

Oke, P.R., Brassington, G.B., Griffin, D.A., Schiller, A.: The Bluelink Ocean Data Assimilation System (BODAS). Ocean Modelling, **21**, 46–70 (2008)

Oke, P. R., Balmaseda, M., Benkiran, M., Cummings, J.A., Dombrowsky, E., Fujii, Y., Guinehut, S., Larnicol, G., Le Traon, P.-Y., Martin M.J.: Observing System Evaluations using GODAE systems, Oceanography, **22**(3), 144-153 (2009)

Oke, P.R., Balmaseda, M., Benkiran, M., Cummings, J.A., Dombrowsky, E., Fujii, Y., Guinehut, S., Larnicol, G., Le Traon, P.-Y., Martin, M.J.: Observational requirements of GODAE Systems. In Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society (Vol. 2), Venice, Italy, 21-25 September 2009, Hall, J., Harrison, D. E. & Stammar, D., Eds., ESA Publication WPP-306, submitted (2010)

Palmer, T.N., Gelaro, R., Barkmeijer, J., Buizza, R.: Singular vectors, metrics, and adaptive observations, J. Atmos. Sci., **55**, 633– 653 (1998)

Rabier, F., Courtier, P., Pailleuz, J., Hollingsworth, A.: Sensitivity of forecast errors to initial conditions. Quart. J. Roy. Meteor. Soc., **122**, 121–150 (1996)

Rabier, F., Gauthier, P., Cardinali, C., Langland, R., Tsyrulnikov, M., Lorenc, A., Steinle, P., Gelaro, R., Koizumi, K.: An update on THORPEX-related research in data assimilation and observing strategies. Nonlin. Proc. Geophys. **15**, 81-94 (2008)

Rao, S.A., Behera, S.K.: Subsurface influence on SST in the tropical Indian Ocean: Structure and interannual variability. Dyn. Atmos. Oceans, **39,** 103–135 (2005)

Rixen, M., Ferreira-Coelho, E.: Operational surface drift prediction using linear and non-linear hyper-ensemble statistics on atmospheric and ocean models. J. Mar. Syst., **65**, 105-121 (2007).

Sakov, P., Oke, P.R.: Objective array design: Application to the tropical Indian Ocean. J. Atmos. Ocean. Tech., **25**, 794-807 (2008)

Schiller, A., Wijffels, S.E., Meyers, G.A.: Design requirements for an Argo float array in the Indian Ocean inferred from observing system simulation experiments. J. Atmos. Ocean. Tech., **21**, 1598–1620 (2004)Schott, F.A., McCreary, J.P.: The monsoon circulation of the Indian Ocean. Prog. Oceanogr., **51**, 1-123 (2001)

Schouten, W.P., de Ruijter, M., van Leeuwen, P.J., Dijkstra, H.A.: An oceanic teleconnection between the equatorial and southern Indian Ocean. Geophys. Res. Lett., **29,** 1812, (2002). doi: 10.1029/2001GL014542

Snyder, C., Joly, A.: Development of perturbations within a growing baroclinic wave. Quart. J. Roy. Meteor. Soc., **124**, 1961-1983 (1998)

Tippett, M.K., Anderson, J.L., Bishop, C.H., Hamill, T.M., Whitaker, J.S.: Ensemble square root filters. Mon. Wea. Rev., **131,** 1485–1490 (2003)

Toth, Z., Kalnay, E.: Ensemble forecasting at NCEP and the breeding method. Mon. Wea. Rev. **125**, 32973319 (1997)

Tracton, M., Kalnay, E.: Operational ensemble prediction at national meteorological center: Practical aspects. Wea. Forecasting, **8**, 379–398 (1993)

Tremolet, Y.: Computation of observation sensitivity and observation impact in incremental variational data assimilation, Tellus, **60**, 964 – 978 (2008)

Vecchi, G.A., Harrison, M.J.: An observing system simulation experiment for the Indian Ocean. J. Climate, **20,** 3300–3319 (2007)

Vidard, A., Anderson, D.L.T., Balmaseda, M.: Impact of ocean observation systems on ocean analysis and seasonal forecasts. Mon. Weath. Rev., **135**, 409-429 (2007)

Wang, X., Bishop, C.H.: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. J. Atmos. Sci., **60**, 1140-1158 (2003)

Wei, M., Frederiksen, J.S.: Error growth and dynamical vectors during southern hemisphere blocking. Nonlin. Proc. Geophys. **11**, 99–118 (2004)

Wei, M., Toth, Z., Wobus, R., Zhu, Y., Bishop, C.H., Wang, X.: Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP. Tellus, **58A**, 28–44 (2006)

Wijffels, S.E., Meyers, G.A.: An intersection of oceanic waveguides: Variability in the Indonesian throughflow region. J. Phys. Oceanogr., **34,** 1232–1253 (2004)