The Oceanographic Data Assimilation Problem: Overview, Motivation and Purposes

Paola Malanotte-Rizzoli\textsuperscript{a} and Eli Tziperman\textsuperscript{b}

\textsuperscript{a}Department of Earth, Atmospheric and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02139

\textsuperscript{b}Weizmann Institute of Science
Rehovot, Israel

Abstract

A brief non-technical overview is given of the data assimilation problem in oceanography. First, a historical perspective is presented that illustrates its main motivations and discusses the objectives of combining fully complex ocean general circulation models (OGCM) and oceanographic data. These objectives are divided into three main categories: model improvement, ocean state estimation and ocean/climate forecasting. Forecasting applications vary from global climate change simulations on a time scale of 50-100 years; through decadal and interannual climate variability, such as the El Nino-Southern Oscillation and the Atlantic thermohaline variability; to extended seasonal forecasts and finally to regional forecast of ocean frontal systems on a time scale of a few weeks. Appropriate assimilation methodologies for each class of oceanographic applications are discussed. For each ocean prediction problem on different time/space scales the needs for data assimilation approaches are pointed out where these are still lacking as they might overcome some of the present deficiencies of the related modeling efforts.

1. INTRODUCTION

The terminology “data assimilation” developed in meteorology about 30 years ago as the methodology in which observations are used to improve the forecasting skill of operational meteorological models. In the practice of operational meteorology, all the observations available at prescribed times are “assimilated” into the model by melding them with the model-predicted values of the same variables in order to prepare initial conditions for the forecast model run.

When used in the oceanographic context, the name data assimilation has acquired a much broader meaning, as reflected in the chapters of this book. Under this general denomination a vast body of methodologies is collected, originating not only in meteorology but in solid-earth geophysics inverse theories and in engineering control theories. All of these methods attempt to constrain a dynamical model with the available data. Moreover, the purposes of oceanographic data assimilation are also often very different from the meteorological case, and three main objectives can be distinguished. One goal is to quantitatively use the data in order to improve the ocean model parameterizations of subgrid scale processes, boundary conditions etc. A second goal is to obtain a four-dimensional realization (the spatial description coupled with the time evolution) of the oceanic flow that is simultaneously
consistent with the observational evidence and with the dynamical equations of motion; the resulting realization can be used for detailed process studies. A third major motivation of ocean data assimilation, the closest to the meteorological one, is to provide initial conditions for predictions of the oceanic circulation. Such predictions are needed in very diverse problems and on very different time scales, from 100 years in climate problems, through interannual climate variability and extended seasonal weather forecasting, to a few weeks in regional ocean forecasting.

In this paper we wish to provide a brief and non-technical overview of the various assimilation problems and methodologies used in oceanography as an introduction to the more specific and technical chapters that follow. Our main focus here is the objectives of oceanographic data assimilation, rather than the methodologies used, and we try to concentrate on what still needs to be done rather than on a review of the existing body of work. Here, as well as in the following chapters, attention is limited to the use of oceanographic data with the most realistic and sophisticated tools presently available to simulate oceanic flows, the ocean general circulation models (OGCM), where one assumes the future of oceanographic data assimilation must lie.

There are many detailed technical references for the various assimilation methodologies used in oceanography, some of them we would like to list here for the reader interested in more technical background information. At the most fundamental levels, inverse methods in oceanography are rather similar to those used in geophysics. Some comprehensive textbooks for this mature field are "Geophysical Data Analysis: Discrete Inverse Theory" by Menke (1984) and "Inverse Problem Theory" by Tarantola (1987). However, these reviews do not meet the requirement of oceanography, that is an analysis of these methods for their application to nonlinear, time-dependent dynamical models of the three-dimensional ocean circulation. From the point of view of the complexity of the physical systems, and of the associated dynamical models, the analysis and application of these methods discussed in Daley's (1991) book, "Atmospheric Data Assimilation", is perhaps the most relevant for oceanographers.

Two major differences still prevent the simple "borrowing" of techniques from meteorology. The first is the motivation for oceanic data assimilation which, as discussed further in the next section, is not as narrowly focused towards short term prediction as are most meteorological efforts. Although it must be added that the motivation for ocean forecasting is rapidly emerging as legitimate and important per se. This book in fact provides important examples of oceanographic operational forecasting. The second reason resides in the major difference between the meteorological and the oceanographic data sets, as further discussed in the next section. This implies that these methodologies, far from being blindly applied to oceanic dynamical problems, must be revisited and sometimes profoundly modified to make them feasible and successful for physical oceanography.

Recent reviews and synthesis of data assimilation methods for oceanographic applications can be found, for example, in the lectures by Miller (1987); in the special issue of Dynamics of Atmospheres and Oceans devoted to Oceanographic Data Assimilation, Haidvogel and Robinson, eds. (1989); and in the review paper by Ghil and Malanotte-Rizzoli (1991). The latter one provides also a very comprehensive review of the literature up to the early 90's. A very recent, thorough synthesis of oceanographic assimilation methodologies is given in Bennett (1992).

2. HISTORICAL PERSPECTIVE

Over the past 25 years or so, since the initial efforts to develop three dimensional ocean circulation models (Bryan, 1969), ocean modeling has made a very significant progress. Chapter 2.1 by Holland and Capotondi provides a review of the milestones in the development and advancement of OGCM's, up to the complexity and sophistication of the present generation of models, capable of most realistic simulations on the global scale.
Chapter 2.1 also offers a perspective of the future possibilities and trends of ocean modeling. In parallel, oceanic observational techniques have been thoroughly revolutionized. However, the lack of a single focusing motivation of oceanic data assimilation such as provided by the need for Numerical Weather Prediction (NWP) in meteorology, caused oceanic models and observational techniques to develop quite independently from each other. When oceanic models and observations started converging, it happened in different paths, depending on the specific objective of each effort.

The early days of oceanography saw dynamic calculations as the main quantitative tool to combine data (temperature and salinity) with "models" (the thermal wind relations). From this modest beginning, relying on highly simplified models and on no formal assimilation procedure, the next step was to introduce a formal least square inverse methodology imported from solid earth geophysics and add the tracer conservation constraints in order to solve the problem of the level of no motion (Wunsch, 1978; Wunsch and Grant, 1982; Wunsch, 1989a,b). This was done in the framework of coarse resolution box models whose dynamics was still very simple although the inverse methodology used was very general. Much of the work done at present on the combination of OGCMs and data stems from the experience obtained in the pioneering work on oceanographic box inverse models.

At the other extreme of model complexity versus sophistication of the assimilation method, efforts began with the "diagnostic models" in which temperature and salinity data were simply inserted into the dynamical equations of fairly complex ocean models in order to evaluate the velocity field (Holland and Hirschman, 1972). The results were very poor due to model-data-topography inconsistencies, and at the next stage, a very simple assimilation methodology was introduced into OGCMs and became known in the oceanographic context as the "robust diagnostic" approach (Sarmiento and Bryan, 1982). The same approach had already been introduced earlier in meteorology as the "nudging" technique (Anthes, 1974) and the term "nudging" has by now become commonly used also in oceanography. In this approach there is no effort to introduce least-square optimality, and the data are just used to nudge the model solution towards the observations at each time step through a relaxation term added to the model equations. The result is far superior to simple diagnostic models, but leaves much to be desired due to the inability to use information about data uncertainty or to estimate the errors in the solution obtained (Holland and Malanotte-Rizzoli, 1989; Capotondi et al., 1995a,b; Malanotte-Rizzoli and Young, 1995).

As the objectives of modeling and observational oceanography began to converge, more formal least square methods taken from meteorology were also used in ocean models, in particular the Optimal Interpolation (OI) method (Robinson et al., 1989; Derber and Rosati, 1989; Mello and Ezer, 1991). OI may be viewed as a nudging technique in which the amount of nudging of the model solution towards the observations depends on the data errors, while also allowing to make error estimates for the solution. This approach, developed in meteorology for NWP, is not capable of improving model parameters or parameterizations, nor is it capable of fitting the entire four dimensional distribution of observations simultaneously to the model solution. However, due to the relatively low computational cost of OI, it is appropriate for higher resolution, short term prediction and state estimation purposes.

Carrying the least square approach for a time dependent model to its rigorous limit, leads to the "Kalman filter/smoothers" assimilation methodology, which is capable of assimilating data into a time dependent model while assuring least-square optimality, full use of a priori error estimates, and calculation of the covariance error matrix for the model outputs. Apart from the fact that the Kalman filter is a formally optimal technique in the least-square sense only for linear models, its high computational cost limits its use at present to simple models, or very coarse OGCMs. Recent efforts are directed at developing efficient even though sub optimal variants of the Kalman filter that allow the use of a full nonlinear OGCM with this method (e.g. Fukumori and Malanotte-Rizzoli, 1995).

The ultimate goal of combining a formal least-square optimization approach with a full complexity OGCM requires the simultaneous solution of hundreds of thousands of coupled
nonlinear equations (the model equations at all grid points and all time steps), and therefore requires an efficient approach which can be found in the "optimal control" engineering literature. This approach, also known as the "adjoint method", is capable of model improvement, parameter estimation and true four dimensional data assimilation. It is equivalent in principle to the Kalman filter (Ghil and Malanotte-Rizzoli, 1991), except that it allows to give up the use and calculation of full covariance matrices, and therefore is more computationally feasible for higher resolution nonlinear OGCMs (Tziperman and Thacker, 1989; Tziperman et al., 1992a,b; Marotzke, 1992; Marotzke and Wunsch, 1993; Bergamasco et al., 1993). The covariance information may be added to the calculation if the computational cost can be afforded.

The development of assimilation methods in physical oceanography seemed to always trail behind meteorology by a few years. This lag is in spite of the fact that the ocean and atmosphere, even though characterized by some important differences, are at the same time similar enough that they can be treated with the same theoretical approaches and methodologies. It is important, therefore, for the ocean modeller to try and understand the reason for this difference in rate of development of data assimilation methodologies in order to be able to isolate potential obstacles for their future use in oceanography.

Clearly a primary reason for the delayed development of oceanic data assimilation was the lack of urgent and obvious motivation such as the need of forecasting the weather and of producing better and longer forecasts as necessary in meteorology. This situation has been rapidly changing in recent years as further discussed in the following section, and ample motivation for ocean data assimilation now exists due to the need for systematic model improvement and for ocean state estimation. The need for ocean prediction is also arising now on various temporal and spatial scales, from climate change predictions, through regional forecasts of the large scale ocean climate variability, e.g. of the North Atlantic thermohaline circulation or El Nino in the Pacific Ocean, to a few weeks regional mesoscale ocean forecasts in frontal regions such as the Gulf Stream system that are required for example by various Naval applications.

The most profound limitation on the development of oceanic data assimilation may have been, however, the lack of adequate data sets. The number of available oceanographic observations is far smaller than the number of meteorological observations, especially when the different temporal and spatial scales are considered. It is estimated, in fact, that the number of presently available oceanographic observations is smaller than its meteorological counterpart by several orders of magnitude (Ghil and Malanotte-Rizzoli, 1991).

New oceanographic data sets, nearly comparable to the meteorological one, i.e. synoptic and with global coverage, are however becoming available. This oceanographic observational revolution of the 90's has been made possible by the advent of satellite oceanography. Already ~40,000 sea surface temperatures are now available daily on a global scale, measured by the NOAA satellites that have been flying since the 80's. In addition, two satellite altimeters are now providing observations of the ocean surface topography that is tightly coupled to ocean currents. The first is TOPEX/POSEIDON, launched in 1992, that is currently producing global maps of sea surface height with a horizontal resolution of ~300 km x 300 km at mid-latitudes every 10 days, and at an impressive accuracy of 5cm (Wunsch, 1994; Fu, 1994; Stammer and Wunsch, 1994). The European satellite ERS-1 is also measuring sea surface topography with higher spatial resolution that resolves the mesoscale eddy field. It also measures the surface wind field on the global scale, at a 1 degree resolution, hence providing information about a crucial driving force for the oceanic circulation. Chapter 2.3 by Fu and Fukumori gives a review of the effects of errors in satellite altimetry for constraining OGCM's through data assimilation. In order to be able to use the altimetric data to study the large scale oceanic circulation, it is however necessary to filter out the effects of tides on the altimetric measurements. The evaluation of global ocean tides can be formulated as an inverse problem and Chapter 3.2 by Egbert and Bennett discusses the possible data assimilation methods.
It is worthwhile to mention two other novel sources of oceanic observations that should help the development of oceanographic data assimilation. The first is the relatively new observational technique of ocean acoustic tomography. Tomography exploits the fact that the ocean is transparent to sound waves and, like in the medical application, the tomographic technique scans the ocean through two-dimensional (vertical or horizontal) slices via sound waves. The difference and novelty of ocean tomography with respect to more traditional point-wise oceanographic measurements lies in the integral nature of the tomographic datum (Worcester et al., 1991). The implications and needs for the assimilation of such integral data into OGCM's are discussed by Cornuelle and Worcester in Chapter 2.4.

A second worldwide major source of oceanographic observations is the World Ocean Circulation Experiment (WOCE) that, through basin wide hydrographic sections, meridional and zonal, should provide a zero-order picture of the large scale global circulation in the 90's. Because hydrographic sections are not synoptic, and are mostly carried out only once, no data of the time evolution will be available and very large water bodies between adjacent sections still remain void of data. Hence the great importance of numerical models endowed with data assimilation capability to act as dynamical interpolators/extrapolators of the oceanic motions. Clearly ocean models and assimilation methods can make better use of the various new and traditional sources of oceanographic data when reliable error estimates are available. Particularly important is the possibility of obtaining estimates of the non-diagonal terms of the error covariance matrices, for which only the diagonal terms, i.e. the data standard deviations, are usually specified. The efforts to obtain such estimates of the full error covariances of traditional oceanographic datasets are discussed by Hogg in Chapter 2.2.

The above brief discussion of the arising needs for ocean data assimilation and the new data sets that are becoming available indicates that possible obstacles to the development of oceanic data assimilation methods have been overcome. Oceanographic data assimilation should now become a fully developed research field. Hence the timeliness of developing modern oceanographic assimilation methods for the OGCM's and the oceanographic data set of the 90's.

3. OBJECTIVES OF OCEANOGRAPHIC DATA ASSIMILATION

Efforts to combine fully complex OGCMs and oceanographic data may roughly be divided into three main categories: model improvement, study of dynamical processes through state estimation, and, finally, ocean/ climate forecast. Let us now consider these objectives in some detail, as well as the relevant assimilation methodologies for each of them.

Even the highest resolution ocean circulation models cannot resolve all of the dynamically important physical processes in the ocean, from small scale turbulence to basin scale currents. There will always be processes that are not represented directly, but rather are parameterized. These parameterizations are sometimes simple, often complicated, and always quite uncertain both in form and in the value of their tunable parameters. Very often, the uncertainty in these parameterizations is accompanied by an extreme sensitivity of the model results to slight variations in them. An obvious though not unique example is the parameterization of small scale vertical mixing in the ocean interior for which many forms have been proposed, and which drastically affects the strength of the thermohaline circulation and the estimate of meridional heat flux of OGCMs (Bryan, 1987). A few other examples are the parameterizations of mesoscale eddies in coarse ocean models used for climate studies (Boning et al., 1995), of mixed layer dynamics (Mellor and Yamada, 1982), and of deep water formation (Visbeck et al., 1994). Another set of uncertain yet crucial parameters corresponds to the poorly known surface forcing by wind stress, heat fluxes and evaporation and precipitation, all of which are subject to typical uncertainties of 30-50% (Trenberth et al., 1989; Schmitt et al., 1989; Trenberth and Solomon, 1993).

Although observations of most of the above unknown model parameters are not available, and many of these parameters are not even directly measurable, there is a wealth of other
oceanographic data that can be used to estimate the unknown parameters. In fact, a most important goal of oceanographic data assimilation is to use the available data systematically and quantitatively in order to test and improve the various uncertain parameterizations used in OGCMs. It is important to understand that by model improvement we refer to the use of data for the determination of model parameters or parameterizations in a way that will result in better model performance when the model is later run without data assimilation. There are typically thousands of poorly known internal model parameters, such as viscosity/diffusivity coefficients at each model grid point, and many thousands if the surface forcing functions are included at every surface grid point (Tziperman and Thacker, 1989). The estimation of these parameters therefore becomes an extremely complicated nonlinear optimization problem which needs to be carried out using efficient methodologies and powerful computers. An assimilating methodology which seems to have the potential to deal with these estimation problems is the conjugate gradient optimization using the adjoint method to calculate the model sensitivity to its many parameters (Hall and Cacuci, 1983; Thacker and Long, 1988). Due to the extreme nonlinearity and complexity of the problem, it is possible however that gradient based methods will not suffice and will need to be combined with some sort of simulated annealing approach to assist in finding a global optimal solution in a parameter space filled with undesired local solutions (Barth and Wunsch, 1990). The adjoint method, while efficient, still requires a significant computational cost when applied to a full OGCM, and is therefore probably limited at present to medium to low resolution ocean models. The resolution of coupled ocean atmosphere models is also limited due to the high computational cost of running them. It is feasible, therefore, that the adjoint method can be used for improving the ocean component of course coupled ocean atmosphere models. A step in this direction is presented in Chapter 3.1 by Sirk et al., who use the adjoint method with a global primitive equation ocean model of a resolution and geometry similar to that used in several recent coupled ocean-atmosphere model studies.

To demonstrate the above general discussion of model improvement by data assimilation, let us now briefly consider two examples of well known difficulties with ocean models that could potentially benefit from data assimilation methodologies. The first is the very strong artificial upwelling in the mid-latitude North Atlantic, in the region inshore of the Gulf Stream (Toggweiler et al., 1989) and in mid-latitudes either using the GFDL (Geophysical Fluid Dynamics Laboratory) model (Sarmiento, 1986; Sugino et al. and Aoki, 1991; Washington et al., 1993) or using the Hamburg large-scale geostrophic model (Maier-Reimer et al., 1993). Boning et al., (1995) show that this upwelling is concentrated in the western boundary layer, roughly between 30° to 40°N and significantly reduces the amount of deep water carried from the polar formation region toward low latitudes and the equator. This strong upwelling is also responsible for the underestimated meridional heat transport in the subtropical North Atlantic which is reduced by about 50% and is due to the deficiency of the parameterization of tracer transports across the Gulf Stream front through the usual eddy diffusivity coefficient. By improving the mixing parameterization using an isopycnal advection and mixing scheme recently proposed by Gent and McWilliams (1990), Boning et al. are able to obtain very substantial improvements in the southward penetration of the NADW (North Atlantic Deep Water) cell and consequently in the meridional heat transport in the subtropical North Atlantic.

The parameterization used by Boning et al is but one of many possible forms, and one would like to see the work of Boning et al. done in an even more thorough and systematic manner, by putting all possible parameterizations into a model, and letting a systematic data assimilation/inverse procedure choose the parameterization and parameters that result in the best fit to the available data.

A second example concerns the difficulty of high resolution ocean models to reproduce the correct separation point of the Gulf stream from the North American continent. This may be due to insufficient model resolution, yet may also be due to imperfect model parameterizations or poor data of surface boundary forcing (Ezer and Mellor, 1992). It is
foreseeable that an improved set of surface boundary conditions may be found through data assimilation, that may eliminate this model problem.

In both of the above examples, the improvement of internal model parameters and of surface boundary conditions via data assimilation may be complemented by a second data assimilation activity, the "state estimation". In this case, model deficiencies are compensated for by using data to force the model nearer to observations during the model run. Thus the strong upwelling found in most simulations of the North Atlantic circulation in the region inshore of the Gulf Stream that results in the shortcut of the thermohaline circulation may be corrected by running the model in a data assimilation mode, rather than as a purely prognostic model.

Such a calculation has been carried out by Malanotte-Rizzoli and Yu (private communication) using the fully nonlinear, time-dependent GFDL code (Cox 1984) and its adjoint first used by Bergamasco et al. (1993) which has been adapted to the North Atlantic ocean to carry out assimilation studies of North Atlantic climatologies (Yu and Malanotte-Rizzoli, 1995). The model is forced by the Hellermann and Rosenstein winds (1983) and the adjoint calculation provides the steady state optimal estimate of the North Atlantic circulation consistent with the Levitus (1982) climatology of temperature and salinity. The assimilation partially corrects for the deficiencies of the analogous purely prognostic calculation. A more realistic meridional thermohaline cell is obtained that protrudes southward much more significantly with ~2/3 of the production rate of 16 Sverdrups (SV) crossing the equator, more closely to the observational figure of ~14 Sv (Schmitz and McCartney, 1993) than in the prognostic simulation. The strong upwelling at 30°N inshore of the Gulf Stream observed by Boning et al. (1995) is in fact eliminated. On the other side, the horizontal wind-driven circulation of the subtropical gyre reconstructed by the adjoint is still too weak, with a maximum Gulf Stream transport of ~60 Sv compared to the value of ~120 Sv found after detachment from Cape Hatteras when encompassing the Southern Recirculation gyre transport (Hogg, 1992). This is due to the smoothed nature of the Levitus climatology showing a "smearered" Gulf Stream front with a cross-section of ~600 km as compared to the realistic values of 200-300 km (Hall and Fofonoff, 1993).

In the case of the Gulf Stream separation point, altimetric and other data can be used to constrain the model to the right separation point (Mellor and Ezer, 1991; Capotondi et al., 1995a,b), and then the resulting model output may be used to study the dynamical processes acting to maintain this separation point. The improved understanding of the dynamics obtained through such uses of data assimilation should eventually result in improved model formulation and more realistic model results.

In spite of the extensive data sets that are becoming available through the new remote sensing methods and the extensive global observational programs mentioned in section 2, the ocean is still only sparsely observed. Most of the interior water mass, and especially the abyssal ocean, remains unexplored. Hence a second aspect of state estimation is the one in which numerical models are constrained by the data to reproduce the available observations, and act as dynamical extrapolators/interpolators propagating the information to times and regions void of data. An especially important example concerns the use of satellite data. It has been shown that ocean models are indeed able to extrapolate instantaneous surface altimetric observations to correctly deduce eddy motions occurring as deep as the main thermocline, at approximately 1,500m (Capotondi et al., 1995a,b; Ghil and Malanotte-Rizzoli, 1991). Clearly this strengthen the case for both the need for data assimilation developments and for satellite altimetry as a global observational system.

The ocean state dynamically interpolated by data assimilation may serve several important goals. On a global scale, unobservable quantities such as the meridional heat flux and the air-sea exchanges can be continuously monitored from the assimilation output to infer possible changes due to climate trends. The knowledge of the natural variability of these quantities is essential for us to be able to differentiate between natural climate variability and a man-induced climate change. On a more regional scale, the high resolution, eddy resolving interpolation of remote sensing data by the models (Mellor and Ezer, 1991; Capotondi et al.,
1995a,b) provides a four dimensional picture of the eddy field which can then be used to study detailed dynamical processes of eddy-mean flow interaction, equatorial wave dynamics, ring formation and ring/jet interactions in the energetic western boundary currents. Such studies, even though they can be done based on the sparse data alone or on model output alone, will gain considerably when carried out on the "synthetic" oceans obtained through data assimilation in dynamical models. Many of the chapters of this book concern the problem of oceanic state estimation through data assimilation. The global applications of the already mentioned chapters 3.1 and 3.2 are related to the estimate of the steady state global circulation (Sirkes et al., Chapter 3.1) and of global ocean tides (Egbert and Bennett, Chapter 3.2). In the tropical ocean, Chapter 4.2 by Busalacchi illustrates how the unique physics of the low-latitude oceans and the wealth of observational data produced by the Ocean Global Atmosphere program have been a catalyst for tropical ocean data assimilation. Among these tropical ocean assimilations are some of the first applications of the Kalman filter and adjoint methods to actual in situ ocean data. These methodologies and related theoretical considerations are discussed in Chapter 4.1 by Miller and Cane. In the context of regional applications, Chapter 5.1 by Malanotte-Rizzoli et al. discusses the development of an efficient and affordable Kalman filter/smoker for a complex, fully nonlinear Primitive Equation model suitable for studies of nonlinear-jet evolutions, model used for realistic simulations of the Gulf Stream system, albeit until now with only a simple nudging assimilation scheme (Malanotte-Rizzoli and Young, 1995).

The third distinct objective of oceanic data assimilation, i.e. ocean and climate nowcasting and prediction, has not been until recently a subject of interest to mainstream oceanography. At present, however, there are more and more specific oceanographic applications in which prediction is not only timely but necessary. It is convenient to classify the oceanographic prediction problems by their time scale, as each of them requires different methodologies of approach and different data.

The problem of climate change is a prediction problem, and therefore needs to be treated as such. Simulation studies of climate change, on a time scale of 50 to 100 years, due to CO2 increase and the greenhouse effect, have recently begun to use coupled ocean-atmosphere models. A very recent study has extended such coupled models simulations to a multiple century time scale (Manabe and Stouffer, 1994). The inclusion of full ocean models in these studies is obviously a step in the right direction considering the significant effect of the ocean on climate on time scales of decades and longer. The use of coupled models is also an important progress from a few years ago when such studies were based on atmospheric models alone, or coupled to a simple mixed-layer ocean models (Wilson and Mitchell, 1987; Schlesinger and Mitchell, 1987; Wetherald and Manabe, 1988; Washington and Meehl, 1989a), or coupled to a model parameterizing heat transport below the mixed layer as a diffusive process (Hansen et al., 1988). Recent studies using fully coupled atmosphere-ocean GCM’s have taken one of two routes in initializing greenhouse warming simulations.

The first route is to initialize the simulation with steady state solutions of the separate ocean and atmosphere sub models obtained by running the two models separately (Stouffer et al., 1989; Manabe et al., 1991; Cubasch et al., 1992; Manabe and Stouffer, 1994). In this procedure the atmospheric model is spun up to a statistical steady state using prescribed SST climatology, such as the Levitus (1982) analysis. The ocean model is then spun-up using boundary conditions which restore the surface temperature and salinity to a similar climatology. The difference in the diagnosed heat and fresh water fluxes from the separate ocean and atmosphere spin-up runs is used to calculate "flux correction" fields. The two models are then coupled, and the flux correction fields are added to the ocean surface forcing at every time step during the subsequent long coupled integration. This correction, while clearly artificial and often of undesirably large amplitude, prevents the quite substantial drifts of the coupled system from the present climate occurring due to the fact that the ocean steady solution is incompatible with the heat and fresh water fluxes provided by the atmospheric model. The initialization of coupled models with steady ocean solutions that are
obtained by restoring the surface model fields of temperature and salinity to climatological data averaged over the last 40 years or more clearly leaves room for significant improvements. This initialization procedure ignores most of the available data which are data from the ocean interior. In addition, the use of many year averaged surface data sets results in a very artificial smoothing and therefore distortion of many important observed features of the oceanic circulation.

The second approach to greenhouse warming simulations is to initialize the model with the observed ocean climatology averaged over tens of years, normally without applying flux correction to avoid a climate drift of the coupled system (Washington and Meehl, 1989). This approach, while avoiding the artificial flux adjustment procedure, suffers from a serious drawback. It is well known from numerical weather prediction that initializing a forecast with the raw data without any weight given to the model dynamics, leads to severe initial "shocks" of the forecast model while it is adjusting to the initial conditions. Such a violent response may be expected in the climate prediction context as well and may severely affect the model response to the greenhouse signal.

What is needed for the climate prediction problem is an assimilation approach that will initialize the prediction simulation using a blending of the data and model results. The initialization should prevent initial shocks, yet constrain the initial condition using the available four dimensional oceanic data base, without the artificial smoothing resulting from the temporal averaging procedure. Such an initialization may also reduce the need for the artificial flux correction procedure.

For such an initialization, a four dimensional global coverage of the ocean is required, as may be provided by programs such as WOCE. Synoptic eddy resolving ocean data are most probably not necessary, as the models used for climate simulations are at this stage far from being eddy resolving, and a precise mapping of the eddy field is not essential for the dynamics in question, but only an overall knowledge of the eddy statistics. Because climate models are fairly coarse due to the high computational cost of these simulations, the assimilation problem can probably be carried out using the more sophisticated assimilation methods, such as the extended Kalman filter or the variational adjoint method. It is important to note, however, that practically nothing has been done so far to address this assimilation/prediction problem which is clearly of paramount interest and importance.

Another coupled climate problem in which prediction is needed is the decadal climate variability problem in which the ocean plays the major role. There are indications, for example, that variability of the North Atlantic thermohaline circulation affects the northern European climate on time scales of 10 to 30 years (Kushnir, 1994). The resulting climate and weather variability has important implications on atmospheric temperature and precipitation over vast regions, is mostly controlled by oceanic processes, and its prediction is of obvious value. The forecasting of decadal climate variability, like that of the global greenhouse problem, needs to be carried out using coupled ocean-atmosphere models and appropriate data sets and assimilation methodology. The mechanisms of the thermohaline variability are still under investigation, with very diverse explanations offered so far, from strongly nonlinear mechanisms (Weaver et al. 1991) suggested using ocean-only model studies to gentler, possibly linear mechanism, based on coupled ocean-atmosphere model studies (Delworth et al., 1994; Griffies and Tziperman, 1994). As the mechanism of this variability is not yet clear, data assimilation could be used to interpolate the little data that exist for this phenomenon, and perhaps clarify the unresolved dynamical issues. The physical mechanisms of decadal climate variability that results from fluctuations of the thermohaline circulation may have important implications concerning the predictability of this variability. Preliminary efforts to examine the predictability of such decadal climate variability are underway (Griffies and Bryan, 1995), yet practically no work has been carried out so far to address this issue as an assimilation and prediction problem, nor are the appropriate data available at present.

An ocean/ climate forecasting problem which presents a successful example of the application of data assimilation methods to ocean/ climate problems is the occurrence of El
Nino-Southern Oscillations (ENSO) in the Pacific equatorial band every three to six years. The profound global socio-economic consequences of this phenomenon have attracted considerable attention in terms of both pure modeling, data collection, and assimilation/forecasting studies.

Barnett et al. (1988) discussed three different approaches used to successfully predict the occurrences of ENSO. One such forecasting scheme uses statistical models that rely on delayed correlations between various indicators in the Equatorial Pacific and the occurrence of ENSO (Barnett, 1984; Graham et al. 1987). A second scheme uses a linear dynamical ocean model that is driven by the observed winds. In the forecast mode, the winds are assumed to remain constant beyond the last time when observations are available, and the ocean model is integrated ahead for a few months to produce the forecast (Inoue and O'Brien, 1984). The third ENSO forecast scheme uses a simple coupled ocean-atmosphere model with linear beta plane dynamics, and a nonlinear equation for the SST evolution. The model is again initialized by running it with the observed winds, and then is integrated further to obtain the forecast (Cane et al., 1986).

Using these various schemes, ENSO occurrences can now be forecast a year in advance with reasonable accuracy. Yet the existing schemes clearly leave room for improvements. Even models that are used now quite successfully for ENSO prediction (Cane et al., 1986) are still fairly simple, with the background seasonal cycle specified in both the atmosphere and the ocean, with linearized dynamics, and with very simplified atmospheric parameterizations. Improvements are needed in the form of fuller models with more realistic parameterization of the oceanic and atmospheric physics, that can simulate both the mean seasonal state and the interannual variability. In addition, the present forecasting schemes do not make full use of the available data, and rely mostly on the observed winds. Better performance may be achieved using more complete assimilation methodologies that use all the available data, including interior ocean data for the temperature, salinity and currents as demonstrated in the recent work by Ji et al. (1995).

Indeed, work is underway to apply the most advanced models and assimilation schemes to the ENSO prediction problem. Until very recently, simple coupled ocean-atmosphere models seemed to be more successful in ENSO forecasting, and fuller primitive equation models had serious difficulties in simulating, not to mention forecasting, ENSO events. This situation is changing now, and full three dimensional primitive equation ocean models coupled to similar atmospheric models are now catching up with the simpler models. Miyakoda et al (1989), for example, have been using such a PE (Primitive Equation) coupled model together with an OI assimilation method to forecast ENSO events. Another direction in which progress has been made is the development of more advanced assimilation methods such as Kalman filtering for this application. As in other applications discussed above, the ENSO prediction problem requires its own variant of these assimilation methodologies, based on the apparently chaotic character of ENSO dynamics (Burger and Cane, 1994; Burger et al., 1995). Chapter 3.3 by Rossa et al. provides an important example of an oceanic four-dimensional data assimilation system developed on the global scale for use in initializing coupled ocean-atmosphere general circulation models (GCM) and to study interannual variability. The model used is a high resolution global ocean model and special attention is given to the tropical Pacific ocean examining the El Nino signature. Chapter 4.3 by Leetma and Ji also provides an example of an ocean data assimilation system developed as a component of coupled ocean-atmosphere prediction models of the ENSO phenomenon, but only for the tropical Pacific configuration. The assimilation system combines various datasets with ocean model simulations to obtain analyses used for diagnostics and accurate forecast initializations. These improved analyses prove to be essential for increased skill in the forecast of sea surface temperature variations in the tropical Pacific.

On a yet shorter time scale, we find the problem of extended seasonal weather prediction, in which again the ocean plays a crucial role. There are many situations in which a seasonal forecast of the expected amount of precipitation, for example, can have a significant impact on agricultural planning, especially in semi-arid regions, but not only there. The application
of coupled ocean-atmosphere GCMs to this problem is at its infancy, and the obvious need for such work can be expected to result in more efforts in this direction in the near future.

It is interesting to note that all the ocean forecasting problems surveyed so far involve using a coupled ocean-atmosphere model, rather than an ocean-only model. There are, however, situations in which ocean-only models can be utilized for relevant short term assimilation and forecasting studies.

A first example for the ocean component alone is given in Chapter 5.2 by Carnes et al. who discuss an ocean modeling-data assimilation monitoring and prediction system developed for Naval operational use in the North Pacific Ocean. Results are presented from three-months long pseudo-operational tests in the effort to address, among other issues, the problem of extended ocean prediction. A further example of forecasts on a very short time scale is given in Chapter 5.3 by Aiken et al., in which a quasi-operational East Coast Forecast system developed to produce 24-hour forecasts of water levels, and the 3-dimensional fields of currents, temperature and salinity in a coastal domain - 24 hour forecasted and observed fields are compared to improve the basic system itself before implementing it with a data assimilation capability.

Finally, an important example is the interest of navies in ocean frontal systems on a time scale of two to four weeks, such as the prediction of the Gulf Stream front and of its meandering. The operational prediction of such synoptic oceanic motions is therefore a primary objective "per se" and a new professional, the ocean forecaster, is rapidly emerging. Like the east coast forecast system of Chapter 5.3, this application is the closest to the meteorological spirit of real-time assimilation and prediction. It involves real time processing and assimilation of remote sensing data, and the production of timely forecasts of front locations and other eddy features in the ocean. A significant body of work already exists for this purpose, and development of such operational forecasting systems is fairly advanced. See, for instance, the issue of Oceanography, Vol. 5, no. 1, 1992 for a review of such operational forecasting systems in the world ocean, with a general discussion of the Navy Ocean Modeling and Prediction Program (Peloquin, 1993) and the interesting DAMEE-GSR effort in the Gulf Stream System involving the assessment of 4 different models through prediction evaluation experiments (Leese et al., 1992; see also Ezer et al., 1992 and Ezer et al., 1993). Chapter 5.4 by Robinson et al. discusses real-time regional forecasting carried out in different areas of the world-ocean. The use and limitations of this methodology are illustrated with practical examples using both a primitive equation and an open ocean quasi-geostrophic model. The latter one constitutes by itself a flexible and logistically portable open-ocean forecasting system, that has been tested in 11 sites of the world ocean comprising frontal systems. All the tests were real-time forecasts, and for six of them the forecasts were carried out aboard ships (Robinson, 1992).

Finally, Chapter 6 by Lozano et al. presents one of the first interdisciplinary applications in developing an ocean prediction system.

4. CONCLUSIONS

Having considered some of the objectives of ocean data assimilation, it is quite surprising to realize how much work is still required to meet them. Much of the effort presently invested in oceanographic data assimilation is in the development of appropriate methodologies, in preparation to approaching the objectives discussed above. The diverse set of objectives discussed here clearly points out that no single assimilation methodology can address all of the needs. It is more likely that several techniques, such as the Kalman Filter, Adjoint Method and Optimal Interpolation will be the main candidates for addressing the future needs of oceanographic assimilation. Each of these methodologies will be used for the specific goals to which it is best suited.

With ample motivation for the combination of fully complex Ocean General Circulation Models and oceanic data, and with new observational techniques and global observational
programs being developed, further developments in oceanic data assimilation are essential. Clearly the needs in this area surpass the invested efforts at this stage, and a significant growth of this research field is needed and may be expected to occur in the very near future.

5. ACKNOWLEDGEMENTS

This research was carried out with the support of the National Aeronautics and Space Administration, Grant #NAGW-2711 (P. Malanotte-Rizzoli).

6. REFERENCES


Fu, L.-L., 1994: Success of the TOPEX/POSEIDON altimetry mission continues, WOCE Notes, Vol. 6, no. 3.


Yu, L and P. Malanotte-Rizzoli, Analysis of the North Atlantic climatologies through the combined OGCM/Adjoint Approach, to be submitted.