Journal of the Korean Society of Oceanography
Vol. 38, No. 2, pp. 52 — 59 June 2003

Data Assimilation for Oceanographic Application: A Brief Overview

SEON K. PARK
Department of Environmental Science and Engineering, Ewha Womans University, Seoul 120-750, Korea

In this paper, a brief overview on data assimilation is provided in the context of oceanographic application. The ocean data assimilation needs to ingest various types of data such as satellites and floats, thus essentially requires dynamically-consistent assimilation methods. For such purpose, sequential and variational approaches are discussed and compared. The major advantage of the Kalman filter (KF) is that it can forecast error covariances at each time step. However, for models with very large dimension of state vector, the KF is exceedingly expensive and computationally less efficient than four-dimensional variational assimilation (4D-Var). For operational application, simplified 4D-Var schemes as well as ensemble KF may be considered.

Keywords: data assimilation, oceanography, variational methods, 4D-Var, Kalman filter

INTRODUCTION

For given limited amount and potential errors of observations, it is essential to extract the maximum amount of information from available measurements, and to incorporate observations distributed in time by considering time evolution of the dynamical system. Thus, a *prediction system* consists of 1) an assimilation system that ingests observational data and extracts useful information from the incomplete data; and 2) a model that describes concerned phenomena well and carries observational information forward to a given prediction time.

Data assimilation (DA) denotes a process in which observational data are melded together with a numerical prediction model, which describes a dynamical flow (ocean or atmosphere), in order to determine the most accurate and complete states of the flow. Using a numerical model, DA usually generates analysis from observations on a grid domain in a way to achieve dynamical consistency. Accurate and efficient estimation of a complex system such as a coastal ocean can be obtained through DA, i.e., blending of data and dynamics.

There exist several remarkable reviews on this subject. A comprehensive description of methods of data analysis and assimilation is provided for meteorology (Daley, 1991; Daley, 1997; Kalnay, 2003) and for oceanography (Anderson and Willebrand, 1989; Robinson *et al.*, 1998; Fukumori, 2001). A wealth of important papers on current methods for DA appears in special issues of *Dyn. Atmos. Oceans* (vol. 13, No. 3-4, 1989), *J. Meteor. Soc. Japan* (vol. 75, No. 1B, 1997), and *J. Marine Systems* (vol. 6, No. 1-2, 1995; vol. 40-41, 2003). Talagrand (1997) provides an elegant theoretical review on current DA methods. Reviews on different scales of ocean DA are given by De Mey (1997) for mesoscale and by Miller *et al.* (1997) for large scales.

Although various techniques are developed for DA, they can be divided into two categories – *sequential* and *variational*. In the sequential approach, the error variances are estimated based on the data. It includes several methods such as successive corrections (Cressman, 1959; Barnes, 1964), optimal interpolation (OI: Gandin, 1963; Robinson *et al.*, 1989), nudging (or Newtonian relaxation: Hoke and Anthes, 1976; Malanotte-Rizzoli and Holland, 1986), and Kalman filtering (Kalman, 1960). A rigorous discussion of present DA methods with special emphasis on sequential methods can be found in Ghil and Malanotte-Rizzoli (1991), whereas recent advances in sequential estimation are discussed in Ghil (1997).

In the variational approach, a functional (often called a *cost function*) is defined to measure misfits between model solutions and observational data, and then an *optimal* initial state of the model is obtained in the process of minimizing the functional using the model dynamics as constraint. In the four-dimen-

*Corresponding author: spark@ewha.ac.kr

sional variational assimilation (4D-Var), the observational information is usually carried backward in time via the so-called *adjoint* model. A collection of papers on variational methods in geosciences appears in Sasaki (1986). Courtier (1997) provides a theoretical review on variational methods. Most recently, Park and Županski (2003) gave a comprehensive review on application of the 4D-Var to mesoscale weather prediction.

Compared to the atmosphere, the ocean includes many dynamical processes that have more deterministic characteristics. However, assimilation of ocean data is still essential for balancing model errors both in dynamics and physics as well as for representing forcings better. DA, which can optimally extract useful information from observations, is regarded as a promising tool for improving prediction skills and is increasingly used in oceanography. In this study, various aspects of DA are briefly discussed, especially in the context of ocean prediction. It is aimed to introduce some advanced DA systems and to discuss the characteristics of each method through non-mathematical descriptions, but not to make a thorough comparison among them.

Section 2 explains modeling and observing systems for ocean. Section 3 provides discussions on various DA methods and their applications. It also introduces some efficient DA methods for operational consideration. Conclusions are given in section 4.

OCEAN MODELING AND OBSERVING SYSTEMS

Governing equations in a numerical model describing ocean circulation are composed of equations of momentum and conservation of heat, salt and mass, and a state equation for sea water. For large scale circulations, which are constrained by the wind, the thermal fluxes and the coasts, the primitive equations (PEs) with hydrostatic balance are normally used instead of a vertical momentum equation. A complete derivation of the PEs for large scale ocean circulation can be found in Veronis (1973). Comprehensive reviews on the ocean circulations and their modelings at various aspects are provided by Niiler (1992) and Haidvogel and Bryan (1992), respectively, in the context of climate system modeling.

Some structures in the ocean are mesoscale – e.g., oceanic eddies and coastal waves. In the mesoscale, the spatial scale is smaller than the planetary scale for which the Rossby number is small. Thus, mesoscale motions are in general well described by the quasi-geostrophic (QG) equations (De Mey, 1997), in which momentum is more dominant than the thermodynamic exchanges.

The QG model assumes that the currents are reasonably represented by geostrophy (though the surface forcing may not necessarily be geostrophic). It also assumes that the horizontal variation in temperature is small – the QG framework does not predict temperature variations and thermodynamics are represented in a simplified form, though it allows the horizontal temperature gradient in the equation. In this case, salt is usually neglected entirely. The QG equations are computationally efficient because the gravity waves are filtered out, thus allowing a long computational time step.

The coastal ocean includes estuaries and the region between the shoreline and the beginnings of the deep ocean (Robinson *et al.*, 1998); thus numerical models dealing with the coastal oceans need to consider a large range of phenomena and scales in both time and space, which should handle with a variety of physical processes of different scales (from millimeter to thousands of kilometers). A comprehensive description for such processes in a coastal ocean prediction system is referred to Robinson *et al.* (1998). For these complex coastal ocean models, data assimilation will be a challenging problem.

Making measurements of the ocean parameters is difficult, especially in the subsurface. Most ships provide observational data on surface ocean and atmosphere on their routes. It is called the comprehensive ocean-atmosphere data set (COADS; NOAA, cited 2003a) and is usually obtained a few months behind real time. But those data provide information neither in the subsurface ocean nor on the global coverage.

Sea surface temperature (SST) is important in assessing interaction between ocean and atmosphere. Measurements of SST have been made routinely for a long time. In these days, SST can also be measured by satellite using infrared or microwave radiometers. The gridded SST data may be available on a monthly and/or weekly basis over much of the globe using an optimal interpolation technique (Reynolds and Smith, 1994). However, they do not tell much about the structure of the subsurface ocean.

Temperature in the upper few hundred meters of water can be measured by a device called an XBT (eXpendable Bathy Thermograph; NOAA, cited 2003b), which consists of a probe connected through a long thin wire to recording equipment on the ship. As the probe drops through the water it measures temperature which is relayed back along the wire. The probe is assumed to have a certain fall rate, and thus a time record of temperature can be transformed into a depth record of temperature.

The World Ocean Circulation Experiment (WOCE) aimed at surveying a large part of the ocean for the first time (WCRP, cited 2003). The WOCE is a multi-year international experiment to enhance knowledge of the ocean circulation, especially the deeper ocean. Much of the early part of the programme was devoted to a one-time survey, to measure accurately the temperature and salinity of the whole depth of the ocean along selected sections, concentrating on measuring parts of the ocean never previously measured. Satellite altimetry played a big role in the WOCE, and a high-quality satellite mission, TOPEX/POSEIDON (NASA, cited 2003), was conducted to try to observe the spatial variability in the ocean from measurements of the top surface, as well as to determine the mean surface elevation from which more accurate measurements of the circulation field could be obtained. After a field phase in 1990-1998, the WOCE has been in the phase of analysis, interpretation, modeling and synthesis (AIMS) until 2002.

Measuring the velocity field directly is more difficult than measuring the temperature field. Fortunately, in much of the ocean, the temperature field is more important for data assimilation (Moore *et al.*, 1987, Anderson and Moore, 1989). However, velocity data are important for understanding the movement of properties by advective processes. One promising approach is to use drifters. Drifters can be deployed at other depths. The first deeper floats were deployed in the sound channel in the ocean, at a depth of 600–800 m where the sound speed is a minimum, making this region a wave guide for sound waves. More recently floats which do not require a tracking array have been developed. They drift at a preset depth but occasionally come to the surface and transmit their position to a satellite, then re-submerge to their operating depth. The rising-up frequency is about 1 month with an operating depth of 1,000 m, though variable. These floats do not require tracking array, thus can be used in remote places. Although they cannot map the eddy field with a surfacing time of around 1 month, they can provide information on mean currents. Assimilating this type of data has not so far been attempted.

In a recent field program named ARGO, a global array of 3,000 free-drifting profiling floats is being deployed to measure temperature and salinity of the upper 2,000 m of the ocean (Wilson, 2002). As those floats move around the ocean, collected data will span a variety of domains in space and time. Assimilation of such data will be a challenging problem.

Further discussions on platforms, sensors, and sampling methods for ocean observations for interdisciplinary data assimilation appear in Dickey (1991, 2003). Busalacchi (1997) provides a comprehensive review on ocean remote sensing as well as some major international ocean field programs such as the Tropical Ocean Global Atmosphere (TOGA: WCRP, 1985), the WOCE, and the Joint Global Ocean Flux Study (JGOFS: JGOFS, cited 2003). Koblinsky and Smith (2001) reviews the ocean observing systems in the 21th century.

DATA ASSIMILATION METHODS AND APPLICATIONS

Assimilation of oceanic observations unavoidably include various types of data from different observing systems (e.g., satellites, ARGO floats, etc). Such observations have different characteristics in terms of spatial and temporal resolutions. When assimilated into a model, this may lead to dynamic and thermodynamic imbalances among model variables. Thus, a proper DA method to deal with this kind of problem should be considered in the oceanic data assimilation.

One of the simplest methods for four-dimensional data assimilation (4DDA) is *nudging* (e.g., Malanotte-Rizzoli and Holland, 1986), which is based on an empirical approach and inserts observation at the nearest model grid point in space and time. The model states are gradually nudged (or relaxed) toward the observations via artificial relaxation terms added in the model equations. However, the initial conditions obtained through nudging are not guaranteed to be dynamically consistent (Bao and Errico, 1997). That is, without some special treatment, the nudging-generated initial conditions will undergo an initial adjustment which may lead to occurrence of dynamical instabilities and computational modes. Recently, some efforts were made to determine optimal nudging coefficients using variational method (e.g., Stauffer and Bao, 1993). Detailed discussions on other assimilation methods (e.g., direct insertion, optimal interpolation, successive correction methods, etc.), which are not dynamically consistent, are given in some review papers and textbooks such as Robinson *et al.* (1998), Fukumori (2001), and Kalnay (2003).

There exist two approaches for the *dynamically consistent* 4DDA - sequential and variational assimilation (see reviews by Ghil and Malanotte-Rizzoli, 1991; Talagrand, 1997; Robinson *et al.*, 1998; Fukumori, 2001). *Sequential assimilation* is based on the optimal estimation theory (Gelb, 1974). This approach was first formulated for a linear system by Kalman (1960), which came to be known as the Kalman filter (KF), and was first introduced in meteorology by Ghil *et al.* (1981) and in oceanography by Ghil and Malanotte-Rizzoli (1991). The KF is designed to optimally estimate the unknown states of a model from noisy data taken at discrete real-time through a linear, unbiased, and minimum error variance recursive algorithm. The KF consists of an *analysis* step, which uses the latest available observations to correct a first guess of the state vector provided by a *forecast* step, i.e., a model-computed time evolution of the former analysis.

For practical application to nonlinear (more precisely quasi-linear) models, the extended KF (EKF) is developed (Jazwinski, 1970; Gelb, 1974; Evensen, 1992); however, the evolution of its error covariance is based on locally linearized dynamics. In unstable regions the EKF will generate unbounded growth of error variance. For models with very large dimension of state vector (say, K), as in meteorological models, this method is exceedingly expensive because the storage and matrix operations in the EKF are proportional to K^2 and K^3, respectively.

Several order-reduction strategies have been developed to remedy problems related to the computational burden in the EKF. One simple approach is to reduce the dimension of the model state vector (e.g., Miller and Cane, 1989; Fukumori *et al.*, 1993), often with asymptotic assumption of constant error covariance matrix (Fu *et al.*, 1993; Fukumori and Malanotte-Rizzoli, 1995; Fukumori, 1995). In another approach, one might consider deducing the KF working space, e.g., undersampling the computational grid used for the filter description (Fukumori, 1995) or making low-rank approximation of error covariance matrices (Cane *et al.*, 1996). The singular extended evolutive Kalman filter (Pham *et al.*, 1998) is a direct reduction of the EKF, given that the covariance matrices can be approximated by such singular low-rank matrices. In the ensemble Kalman filter (Evensen 1994), which can be used for nonlinear models, the Monte Carlo approximation is made for filtering the error covariance. The last approach relies on physical

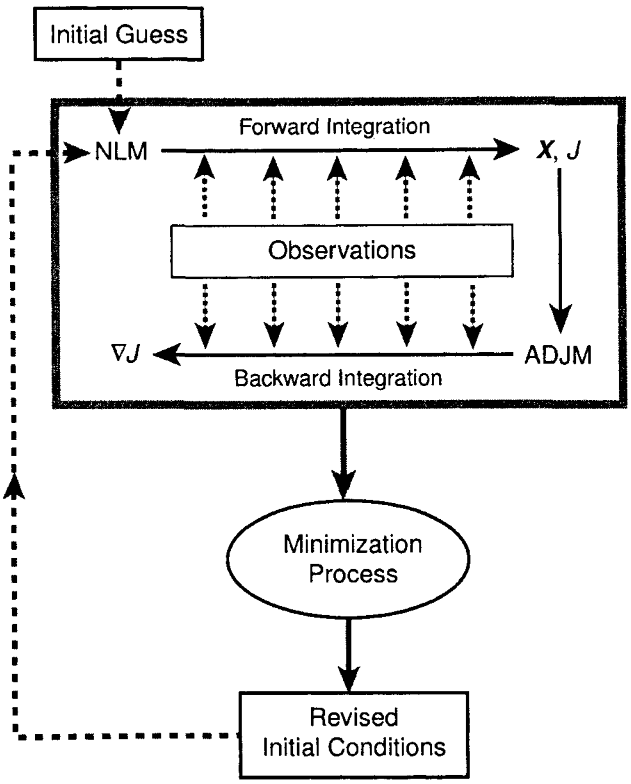

Fig. 1. Processes of 4D-Var in one iteration. From Park and Zupanski (2003).

considerations that simplify the model error dynamics and decrease the degree of freedom of the system – for example, geostrophic approximation (Dee, 1991), long wave approximation in the tropics (Cane and Patton, 1984), etc.

Variational assimilation is based on the optimal control theory (Lions, 1971) and the adjoint formulation (Marchuk, 1975). The four-dimensional variational data assimilation (4D-Var) is developed to obtain optimal states of the atmosphere using multi-time-level observations through model dynamics. The major objective of 4D-Var is to obtain optimal input parameters (e.g., initial conditions, model error, lateral boundary conditions, parameterization coefficients, etc.) by globally fitting model solutions to all available observations over an interval of time (i.e., assimilation period). The 4D-Var involves an iterative process, which normally takes several tens iterations, to minimize a cost function that measures the square distance between the model solutions and the observations, using the model as constraint (i.e., constrained minimization). A backward run of the adjoint model is carried to provide the gradient information into the minimization process. Figure 1 shows the processes involved in one iteration of 4D-Var, which con-

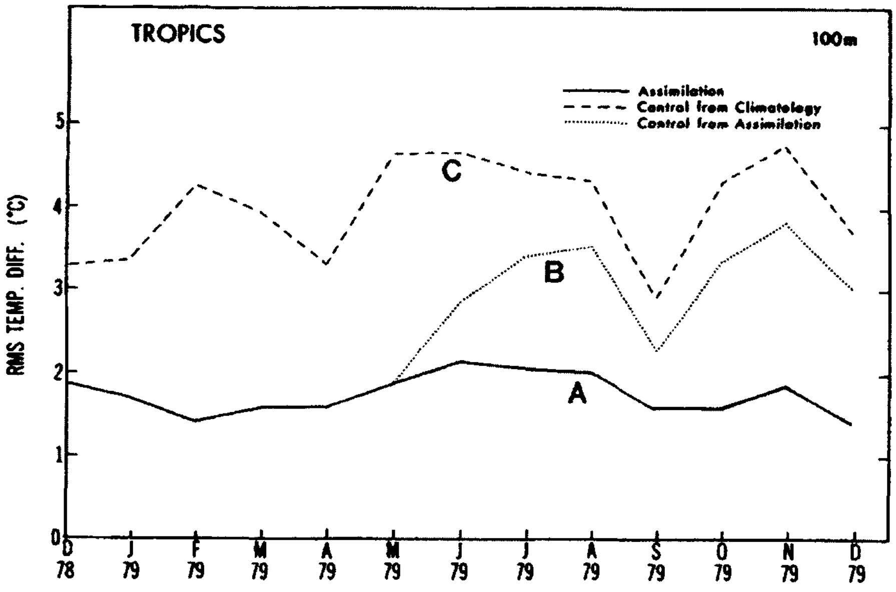

Fig. 2. Effect of data assimilation in the RMS temperature differences. From Derber and Rosati (1989).

sists of one forward nonlinear model run, one backward adjoint model run, and a minimization process.

The 4D-Var approach have been applied to assimilate altimetry data into the QG models (Morrow and De Mey, 1995), the shallow water models (Cong *et al.*, 1998), and the PE models (Lee and Marotzke, 1997). Courtier (1997) provides a review on the theory of 4D-Var, and Park and Županski (2003) on applications to small scale meteorology. A comprehensive review on theory and application to the ocean circulation problems is given by Fukumori (2001).

The 3D-Var method provides a suboptimal solution to the model states using only one-time-level observation, thus less general and computationally less demanding (e.g., Courtier *et al.*, 1998). However, even the 3D-Var method can improve the forecast substantially. Figure 2 depicts an example of applying 3D-Var to temperature forecast in tropical ocean (see Derber and Rosati, 1989). In case A data are assimilated throughout, while in case C no data are assimilated. In case B, data are assimilated only during an early period. When assimilation is switched off in B, it quickly reverts towards the no-assimilation case C. It demonstrates the importance of *continuous* data assimilation.

For a purely linear problem, assuming that the same error covariances are used, the 4D-Var becomes identical to the Kalman smoother (KS - composed of computing a forward KF estimator and a backward estimator for a time interval, and finding optimal linear combination between the two; see Zhu *et al.*, 1999). Li and Navon (2001) reported that various properties of equivalence of the 4D-Var to the KF and KS result from optimality properties of 4DVAR. The KS (or KF) explicitly evolves error covariance functions and can forecast error covariances (so-called cycling) while the 4D-Var only implicitly evolves error covariances during the assimilation period and, if not combined with some other method, cannot forecast and update error covariances from one data assimilation cycle to another.

However, the KF carries observational information only forward in time (i.e., from past to future), while the 4D-Var can carry information both forward and backward in time. It is also found that the 4D-Var is more successful than the EKF in reconstruction of the physical fields in data-void regions (see Ehrendorfer, 1992). Due to the enormous computational demand required by KF, the cheaper 4D-Var is considered to be more appropriate for operational application at the present time. A recently developed technique, the ensemble KF (Evensen, 1994), has shown great potential of saving computing time (e.g., Houtekamer and Mitchell 1998), even making it possible to consider operational applications.

A direct comparison between the EKF and the 4D-Var (but not in the computational viewpoint) is made by Ikeda *et al.* (1995) for assimilating simulated altimeter data into a QG-based linear Rossby wave model. They showed that the EKF was capable of generating the proper vertical structure of the error covariance functions, and the 4D-Var efficiently integrated solution-data mismatch. The major difference between the two methods was that the 4D-Var could reconstruct fields prior to observations, which was not possible with the EKF.

Other advantages of 4DVAR include, as indicated by Park and Županski (2003): 1) ability to accept observations in their 'raw' (or near-raw) format, thus eliminating the need for retrieval operations; 2) linkage of the observations to the model variables in a nonlinear manner; 3) projection of information from the model space to observation space, and vice versa, in a consistent manner through an observation operator; 4) implicit use of flow-dependent structure functions which enables transfer of local information to all model grids.

The conventional cost function used in the 4D-Var includes a term measuring the distance to the background at the beginning of the interval, and a summation over time of each observational increment (see Kalnay, 2003). In this case, the model is assumed to be perfect and is used as a *strong* constraint (Sasaki, 1970a). In a weak constraint 4D-Var, the model errors are accounted for in the cost function (Sasaki, 1970b; Griffith and Nichols, 1996). Županski *et al.* (2002) showed that the 4D-Var results using the forecast model as a weak constraint were superior to those with the strong constraint. Recently, Bennet (1992) developed the so-called representer method, which solves the variational problem with weak constraints by seeking a solution linearly expanded into data influence functions (i.e., representers) that correspond to each separate observation. Then the assimilation problem tries to determine the optimal coefficients of the representers within the observational space rather than the model state space, thus requiring less computing time (e.g., Egbert *et al.*, 1994; Kivman, 1997). An algorithmic overview on this method is referred to Chua and Bennett (2001).

Since the standard 4D-Var still requires a large computational time, some efficient methods have been developed. An incremental 4D-Var is developed by Courtier *et al.* (1994), in which minimization is performed for a coarse domain with simple physics and then the increment is added to the full nonlinear forecast with fine resolution. Huang *et al.* (1997) developed a poor man's 4D-Var by using the increment obtained through the adjoint run into the 3D-OI. In the inverse 3D-Var, a quasi-inverse linear model is used to obtain an increment (Kalnay *et al.*, 2000). The incremental 4D-Var scheme is in operational mode in several weather centers in these days. Although these simplified 4D-Var schemes are suboptimal, as those KF schemes with order reduction, operational applications demonstrate that their results are much better than those obtained from assimilation methods that are not dynamically consistent.

CONCLUSIONS

In this study, a brief overview of data assimilation is provided in the context of oceanic prediction system. Basically the ocean data assimilation requires to ingest various types of data including observations from satellites and ARGO floats. Among many methods for data assimilation, discussions were focused on the dynamically-consistent schemes represented by the Kalman filtering (KF) and the 4D-Var. The KF method demands a huge computer resources for a model with a large dimension of state vector, and several alternative strategies have been developed to reduce the computational burden. Since the 4D-Var is capable of accepting data in their raw format, it can be an appropriate assimilation method to deal with many different types of ocean data. However, to properly handle with such data, proper nonlinear observation operators should be developed.

Although the 4D-Var has many advantages, it still demands a large amount of computing time and resources. In operational sense, simplified 4D-Var methods, such as incremental 4D-Var, poor man's 4D-Var and inverse 3D-Var, and their hybrid application should be considered (see Park and Županski, 2003 for a review). The ensemble KF is also a good candidate for operational application but may need rigorous tests. The model error should also be taken into account as a weak constraint in the operational data assimilation.

An elegant and practical way of developing and testing an ocean assimilation system is to incorporate artificial data into the ocean models. In this case, (assimilated) data can be either taken from the same model at a different point in its integration cycle or from a different, usually more sophisticated, model. Such so-called twin OSSEs (observing-systems simulation experiments; see Arnold and Dey, 1986) allow introduction of limited data sets distributed in space and time to test whether more complete data can be recovered in a situation in which the true circulation is already known in all respects. Thus, before we jump into any DA system with complex ARGO data, it is desirable to perform rigorous tests using the OSSEs strategy. With some ARGO data available, it might be possible to proceed with the OSEs (observing-systems experiments) for more realistic tests of a DA system.

ACKNOWLEDGEMENTS

The author is grateful to an anonymous reviewer who provided detailed comments. This paper is improved significantly by those comments. This study is partly supported by the ARGO Program of the Meteorological Research Institute (METRI), Korea Meteorological Administration (KMA) and by the Meteorological and Earthquake R & D Programs for Improving the Medium-Range Forecasting funded by the KMA. Additional support is from the Agency for Defense Development of Korea through the Fundamental Research program: Research on the data assimilation methods for the real-time forecast model of the East Sea (UD030005AD).

REFERENCES

Anderson, D.L.T. and A.M. Moore, 1989. Initialization of equatorial waves in ocean models. *J. Phys. Oceanogr.*, **19**: 116–121.

Anderson, D.L.T. and J. Willebrand, 1989. Ocean Circulation Models: Combining Data and Dynamics. Kluwer, 605 pp.

Arnold, C.P., Jr. and C.H. Dey, 1986. Observing-systems simulation experiments: Past, present, and future. *Bull. Amer. Meteor, Soc.*, **67**: 687–695.

Bao, J.W. and R.M. Errico, 1997. An adjoint examination of a nudging method for data assimilation. *Mon. Wea. Rev.*, **125**: 1355–1373.

Barnes, S., 1964. A technique for maximizing details in numerical map analysis. *J. Appl. Meteor.*, **3**: 395–409.

Bennett, A.F., 1992. Inverse Methods in Physical Oceanography. Cambridge University Press, 346 pp.

Busalacchi, A.J., 1997. Oceanic observations. *J. Meteor. Soc. Japan*, **75**(1B): 131–154.

Cane, M.A. and R.J. Patton, 1984. A numerical model for low-frequency equatorial dynamics. *J. Phys. Oceanogr.*, **14**: 1853–1863.

Cane, M.A., A. Kaplan, R.N. Miller, B. Tang, E.C. Hackert and A.J. Busalacchi, 1996. Mapping tropical Pacific sea level: data assimilation via a reduced state Kalman filter. *J. Geophys. Res.*, **101**(C10): 22599–22617.

Chua, B.S. and A.F. Bennett, 2001. An inverse ocean modeling system. *Ocean Modeling*, **3**: 137–165.

Cong, L.Z., M. Ikeda and R.M. Hendry, 1998. Variational assimilation of Geosat altimeter data into a two-layer quasi-geostrophic model over the Newfoundland ridge and basin. *J. Geophys. Res.*, **103**(C4): 7719–7734.

Courtier, P., 1997. Variational methods. *J. Meteor. Soc. Japan*, **75**(1B): 211–218.

Courtier, P., J.-N. Thèpaut and A. Hollingsworth, 1994. A strategy for operational implementation of 4-DVAR using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**: 1367–1387.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier and M. Fisher, 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation. *Quart. J. Roy. Meteor. Soc.*, **124**: 1783–1808.

Cressman, G.P., 1959. An operational objective analysis system. *Mon. Wea. Rev.*, **87**: 367–374.

Daley, R., 1991. Atmospheric Data Analysis. Cambridge University Press, 457 pp.

Daley, R., 1997. Atmospheric data assimilation. *J. Meteor. Soc. Japan*, **75**(1B): 319–329.

Dee, D.P., 1991. Simplification of the Kalman filter for meteorological data assimilation. *Quart. J. Roy. Maeteor. Soc.*, **117**: 365–384.

De Mey, P., 1997. Data assimilation at the oceanic mesoscale: A review. *J. Meteor. Soc. Japan*, **75**(1B): 415–427.

Derber, J.C. and A. Rosati, 1989. A global data assimilation system. *J. Phys. Oceanogr.*, **19**: 1333–1347.

Dickey, T.D., 1991. The emergence of concurrent high resolution physical and bio-optical measurements in the upper ocean and their applications. *Rev. Geophys.*, **29**: 383–413.

Dickey, T.D., 2003. Emerging ocean observations for interdisciplinary data assimilation systems. *J. Mar. Systems*, 40–41: 5–48.

Egbert, G.D., A.F. Bennett and M.G.G. Foreman, 1994. TOPEX/POSEIDON tides estimated using a global inverse model. *J. Geophys. Res.*, **99**(C12): 24821–24852.

Ehrendorfer, M., 1992. Four-dimensional data assimilation: comparison of variational and sequential algorithms. *Quart. J. Roy. Meteor. Soc.*, **118**: 673–713.

Evensen, G., 1992. Using the extended Kalman filter with a multilayer quasigeostrophic ocean model. *J. Geophys. Res.*, **97**(C11): 17905–17924.

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**(C5): 10143–10162.

Fu, L.L., I. Fukumori and R.N. Miller, 1993. Fitting dynamic models to Geosat sea-level observations in the tropical Pacific ocean: II. A linear, wind-driven model. *J. Phys. Oceanogr.*, **23**: 2162–2181.

Fukumori, I., 1995. Assimilation of Topex sea level measurements with a reducedgravity, shallow water model of the tropical Pacific ocean. *J. Geophys. Res.*, **100**(C12): 25027–25039.

Fukumori, I., 2001. Data assimilation by models. In: Satellite Altimetry and Earth Sciences, edited by Fu, L.-L. and A. Cazenave, Academic Press, 237–265.

Fukumori, I. and P. Malanotte-Rizzoli, 1995. An approximate Kalman filter for ocean data assimilation: an example with an idealized Gulf stream model. *J. Geophys. Res.*, **100**(C4): 6777–6793.

Fukumori, I., J. Benveniste, C. Wunsch and D.B. Haidvogel, 1993. Assimilation of sea surface topography into an ocean circulation model using a steady state smoother. *J. Phys. Oceanogr.*, **23**: 1831–1855.

Gandin, L.S., 1963. Objective Analysis of Meteorological Fields. GIMIZ, Leningrad. English translation by Israeli Program for Scientific Translation, Jerusalem, 1965, 242 pp.

Gelb, A., Ed., 1974. Applied Optimal Estimation. The MIT Press, 374 pp.

Ghil, M., 1997. Advances in sequential estimation for atmosphere and oceanic flow. *J. Meteor. Soc. Japan*, **75**(1B): 289–304.

Ghil, M. and P. Malanotte-Rizzoli, 1991. Data assimilation in meteorology and oceanography. *Adv. Geophys.*, **33**: 141–266.

Ghil, M., S. Cohn, J. Tavantzis, K. Bube and E. Isaacson, 1981. Applications of estimation theory to numerical weather prediction. In: Dynamic Meteorology: Data Assimilation Methods, edited by Bengtsson, L., M. Ghil and E. Källén, Springer-Verlag, 139–224.

Griffith, A.K. and N.K. Nichols, 1996. Accounting for model error in data assimilation using adjoint models. In: Computational Differentiation: Techniques, Applications, and Tools, SIAM, pp 195–204.

Haidvogel, D.B. and F.O. Bryan, 1992. Ocean general circulation modeling. In: Climate System Modeling, edited by Trenberth, K.E., Cambridge University Press, 371–412.

Hoke, J. and R. Anthes, 1976. The initialization of numerical models by a dynamic relaxation technique. *Mon. Wea. Rev.*, **104**: 1551–1556.

Houtekamer, P.L. and H.L. Mitchell, 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**: 796–811.

Huang, X.-Y., A. Cederskov and Källén, E., 1997. Using an adjoint model to improve an optimum interpolation-based data assimilation system. *Tellus*, **49A**: 161–176.

Ikeda, M., G. Evensen and L. Cong, 1995. Comparison of sequential updating, Kalman filter and variational methods for assimilating Rossby waves in the simulated Geosat altimeter data into a quasi-geostrophic model. *J. Mar. Systems*, **6**: 15–30.

Jazwinski, A.H., 1970. Stochastic Processes and Filtering Theory. Academic Press, 376 pp.

JGOFS, cited 2003. The Joint Global Ocean Flux Study. Available online from "http://www.uib.no/jgofs/jgofs.html".

Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *Trans. ASME, Series D, J. Basic Eng.*, **82**: 35–45.

Kalnay, E., 2003. Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, 341 pp.

Kalnay, E., S.K. Park, Z.-X. Pu and J. Gao, 2000. Application of the quasi-inverse method to data assimilation. *Mon. Wea. Rev.*, **128**: 864–875.

Kivman, G.A., 1997. Weak constraint data assimilation for tides in the Arctic Ocean. *Prog. Oceanogr.*, **40**: 179–196.

Koblinsky, C.J. and N.R. Smith, Eds., 2001. Observing the Oceans in the 21st Century. GODAE, Bureau of Meteorology, Melbourne, Australia, 580 pp.

Lee, T. and J. Marotzke, 1997. Inferring meridional mass and heat transports of the Indian-Ocean by fitting a general-circulation model to climatological data. *J. Geophys. Res.*, **102**(C5): 10585–10602.

Li, Z. and I.M. Navon, 2001. Optimality of variational data assimilation and its relationship with the Kalman filter and smoother. *Quart. J. Roy. Meteor. Soc.*, **127**: 661–683.

Lions, J.L., 1971. Optimal Control of Systems Governed by Partial Differential Equations. Springer-Verlag, 396 pp.

Malanotte-Rizzoli, P. and W.R. Holland, 1986. Data constraints applied to models of the ocean general circulation. I. The steady case. *J. Phys. Oceanogr.*, **16**: 1665–1682.

Marchuk, G.I., 1975. Formulation of theory of perturbations for complicated models. *Appl. Math. Optimization*, **2**: 1–33.

Miller, R.N. and M. Cane, 1989. A Kalman filter analysis of sea level height in the tropical Pacific. *J. Phys. Oceanogr.*, **19**: 773–790.

Miller, R.N., A.J. Busalacchi and E.C. Hackert, 1997. Application of data assimilation to analysis of the ocean on large scales. *J. Meteor. Soc. Japan*, **75**(1B): 445–462.

Moore, A.M., N.S. Cooper and D.L.T. Anderson, 1987. Data assimilation in models of the Indian Ocean. *J. Phys. Oceanogr.*, **17**: 1965–1977.

Morrow, R. and P. De Mey, 1995. Adjoint assimilation of altimetric, surface drifter and hydrographic data in a QG model of the Azores Current. *J. Geophys. Res.*, **100**(C12): 25007–25025.

NASA, cited 2003. Ocean Surface Topography from Space. Available online from "\http://topex-www.jpl.nasa.gov/mission/topex.html".

Niiler, P.P., 1992. The ocean circulation. In: Climate System Modeling, edited by Trenberth, K.E., Cambridge University Press, 117–148.

NOAA, cited 2003a. COADS: The Comprehensive Ocean-Atmosphere Data Set Project. Available online from "\http://www.ncdc.noaa.gov/oa/climate/coads/".

NOAA, cited 2003b. The Expendable Bathythermograph (XBT). Available online from "\http://www.aoml.noaa.gov/phod/uot/uot_xbt.html".

Park, S.K. and D. Županski, 2003. Four-dimensional variational data Assimilation for mesoscale and storm-scale applications. *Meteor. Atmos. Phys.*, **82**: 173–208.

Pham, D.T., J. Verron and M.C. Roubaud, 1998. A singular evolutive extended Kalman filter for data assimilation in oceanography. *J. Mar. Systems*, **16**: 323–340.

Reynolds, R.W. and T.M. Smith, 1994. Improved global sea surface temperature analysis using optimum interpolation. *J. Climate*, **7**: 929–948.

Robinson, A.R., M.A. Spall, L.J. Walstad, and W.G. Leslie, 1989. Data assimilation and dynamical interpolation in Gulfcast experiments. *Dyn. Atmos. Oceans*, **13**: 301–316.

Robinson, A.R., P.F.J. Lernusiaux and N.Q. Sloan III, 1998. Data assimilation. In: The Sea, edited by Brink, K.H. and A.R. Robinson, pp. 541–594.

Sasaki, Y.K., 1970a. Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.*, **98**: 875–883.

Sasaki, Y.K., 1970b. Numerical variational analysis with weak constraint and application to surface analysis of severe storm gust. *Mon. Wea. Rev.*, **98**: 899–910.

Sasaki, Y.K., 1986. Variational Methods in Geosciences. Elsevier, 309 pp.

Stauffer, D.R. and J.-W. Bao, 1993. Optimal determination of nudging coefficient using the adjoint equations. *Tellus*, **45A**: 358–369.

Talagrand, O., 1997. Assimilation of observations, an introduction. *J. Meteor. Soc. Japan*, **75**(1B): 191–209.

Veronis, G., 1973. Large scale ocean circulation. *Adv. Appl. Mech.*, **13**, 1–92.

Wilson, S., 2002. Launching the Argo Armada: Taking the ocean's pulse with 3,000 free-ranging floats. *Oceanus*, **42**: 17–19.

WCRP, 1985: Scientific Plan for the Tropical Ocean and Global Atmosphere Programme. WCRP Publication Series, No. 3. World Meteorological Association, Geneva, 146+xxvii pp.

WCRP, cited 2003. The World Ocean Circulation Experiment (WOCE) 1990-2002. Available online from "\http://www.soc.soton.ac.uk/OTHERS/woceipo/ipo.html".

Zhu, Y., R. Todling and S. E. Cohn, 1999. Technical remarks on smoother algorithms. DAO Office Note 99-02, NASA/GSFC, 45 pp.

Županski, D, M. Županski, E. Rogers, D.F. Parrish and G.J. DiMego, 2002. Fineresolution 4DVAR data assimilation for the Great Plains tornado outbreak of 3 May 1999. *Wea. Forecasting*, **17**: 506–525.

Manuscript received February 27, 2003
Revision accepted May 28, 2003
Editorial handling: Jae-Hak Lee