Application of data assimilation to ocean and climate prediction

1 Introduction

Variations in ocean heat content account for much of the energy that drives weather and climate systems; errors in the representation of the ocean dynamics and thermodynamics in computational models affect the validity of forecasts of the ocean and atmosphere on daily, seasonal and decadal time scales. To improve the forecasts, a state estimation technique from control theory, known as 'data assimilation' [4], is used to combine current estimates of state variables such as temperature, salinity and velocity from a numerical model with observations of the ocean. The data assimilation process generally works well with ocean models away from the equator, but spurious ocean circulation occurs when thermal data are assimilated near the equator into models with systematic errors. Research undertaken by the University of Reading and the Met Office investigated these errors and developed a novel pressure correction technique that improves the analyses and forecasts by restoring dynamical balances, eliminating spurious deep overturning circulations [2]. The technique has been implemented by the Met Office and by the European Centre for Medium Range Weather Forecasting (ECMWF) in their forecasting systems, resulting in major improvements to ocean forecasting and seasonal prediction using coupled ocean-atmosphere models.

2 Data Assimilation

A computational model can never completely describe the complex physical processes involved in the behaviour of a real world dynamical system. Data assimilation techniques are used to improve predictions from the numerical models. Mathematically, the assimilation problem is an inverse problem, matching model data to observations [4]. To illustrate the problem we assume we have a model of the evolution of the height $h(x;y;t)$ of the ocean surface and that at a given time t we have observations of the surface height at N locations $(x_n;y_n); n = 1;2;:::;N;$ which we write in a vector $\mathbf{h^o}$ of dimension $N:$ We suppose also that we have a prior forecast (or background) vector of surface heights at these same points, $\mathbf{h^b}$, generated by the model dynamics. An improved estimate of the heights is then obtained by minimizing the differences between the measurements and the prior model forecast states, weighted by forecast and measurement error statistics. Statistically, the Best Linear Unbiased Estimate (BLUE) of $\mathbf{h};$ is obtained by minimizing the objective function

$$J(\mathbf{h}) = (\mathbf{h} - \mathbf{h^o})^{\mathbf{T}}\mathbf{R}^{-\mathbf{1}}(\mathbf{h} - \mathbf{h^o}) + (\mathbf{h} - \mathbf{h^b})^{\mathbf{T}}\mathbf{B}^{-\mathbf{1}}(\mathbf{h} - \mathbf{h^b}): \qquad (1)$$

with respect to $\mathbf{h}$, where $\mathbf{R}$ and $\mathbf{B}$ are square, symmetric positive-definite matrices representing the covariances of the errors expected (v)15for1o, 4.114 e50(4.114 (xpected)7220(by)-2Centre)]TJ 0 -11.

dict. It is because of nonlinear transfers between scales within fluids that ocean and weather predictions have an inherently stochastic nature.

The updated vector of heights from the minimization is used to initialize and evolve the numerical model forward to obtain a forecast at the next time that observations are available; the minimization process is then repeated. In practice, if only the surface heights are updated, poor forecasts may be obtained. Ideally we would also update the horizontal velocities u and v that describe the evolution of the flow. The geostrophic balance relationship between the forces per unit mass due to gradients in the surface height fields and the Coriolis acceleration can be used for example, away from the equator, to update the velocities and provide better initial conditions for predictions. Alternatively a dynamical model, such as the shallow water equations, can be used to forecast and assimilate the heights and the horizontal velocities simultaneously.

In practice we do not have as many observations as we have state variables in our model equations and the locations of the observations are not necessarily at the same locations in space as our forecast model states. In this case we map the model states to variables that can be compared directly to the observations, using an ‘observation operator’; this could be an interpolation operator, for example. We then minimize the differences between the measurements and the observation operator applied to the vector of model states, together with the differences between the prior forecast (background) and the model states, weighted by the measurement and forecast error statistics respectively. This leads to a nonlinear least-squares problem with an objective function similar to (1); the minimum of the objective function provides an improved estimate of all the forecast states. The problem is well-posed even if the number of observations is less than the number of state variables. Uniqueness of the solution is ensured because the covariance matrix of the errors in the prior forecast (background) field is nonsingular (see [4]).

3 Unrealistic motions generated by data assimilation near the equator

In addition to improving forecasts from the model, data assimilation can be used to identify systematic errors within the model, highlighting where the model forecasts are consistently incorrect in relation to observations. An analysis of early results obtained from a system assimilating measurements of thermal profiles into a global ocean model revealed that in the eastern equatorial Pacific, where there was a particularly good source of data from the Tropical Atmosphere-Ocean (TAO) mooring array, the data assimilation was continuously making large temperature updates of the same sign. Heating by these assimilation updates was generating “equatorial bonfires” involving large unrealistic vertical velocities and overturning circulations extending to considerable depths (see Figure 1a).

To understand this behaviour we first note that near the equator the Coriolis acceleration is small and flows along the equator are not in geostrophic balance. In the ocean to a first approximation the momentum balance along the equator in the near surface layers is between momentum input by the surface wind stress, which is mixed downward by turbulent motions, and the pressure gradient along the equator. The surface wind stresses and turbulent mixing at the equator are not accurately known and measurements of salinity (which also affects the density and hence the pressures within the ocean) are also limited (although they are much more abundant since the Argo array of profiling floats reached maturity in about 2004). As an ocean

$$\frac{\partial u^t}{\partial t} \quad f v^t = \quad g \frac{\partial h^t}{\partial x} + t_x^t; \tag{2}$$

$$\frac{\partial v^t}{\partial t} + f u^t = \quad g \frac{\partial h^t}{\partial y} + t_y^t; \tag{3}$$

$$\frac{\partial h^t}{\partial t} + H$$

$$\frac{\partial h^e}{\partial t} - g h^e = 0. \qquad (12)$$

As $g \neq 0$, (12) implies that any stationary or time-mean solution has $h^e = 0$, which means that the model surface height is the same as the true surface height and the time-mean data assimilation updates will be zero. Using $h^e = 0$ in (11) we see also that the horizontal divergence $\partial u^e/\partial x + \partial v^e/\partial y$ is zero for these solutions. For the 3D problem this implies that the vertical velocities will also be zero. It is also possible to show that there are no spurious sources of energy in the augmented equation set (5) - (8) and that they do not support spurious wave amplification or propagation.

Initial integrations showed that the scheme reduced the time-mean temperature increments and improved the equatorial undercurrent. The method was implemented in the Met Office ocean assimilation system and shown to have a significant impact on the forecast [2]. In Figure 1a we see the substantial overturning deep ocean vertical velocities induced by the assimilation and in Figure 2a we see the mean temperature updates from the assimilation that generate this flow. In Figures 1b and 2b,

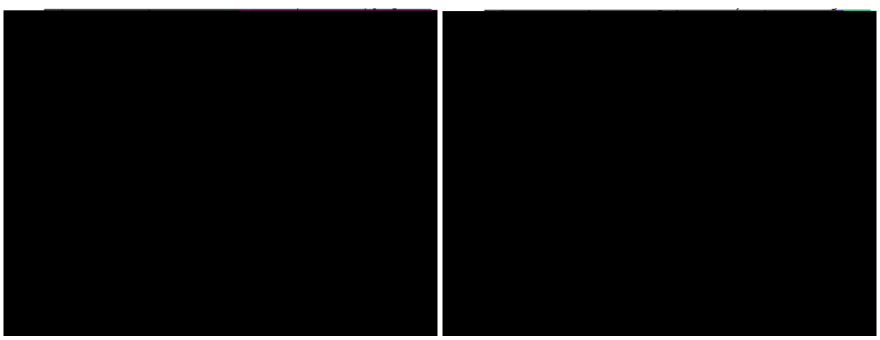

Fig. 1 Annual mean vertical velocities (cm s^{-1}) cross-section across the equator at 110 W: (a) with standard data assimilation and (b) with the pressure correction scheme.

we see that the pressure correction technique provides a major improvement in the assimilation, eliminating the spurious circulation and reducing the temperature updates.

5 Impacts

Ocean data assimilation systems are used to initialise predictions of the ocean and climate from days to decades ahead, as well as providing information about the past through reanalyses. These systems use a vast quantity of input data collected from satellites, surface drifting buoys, sub-surface profiling floats from the Argo array, moored buoys, ships and various other observing platforms. Improvements to data

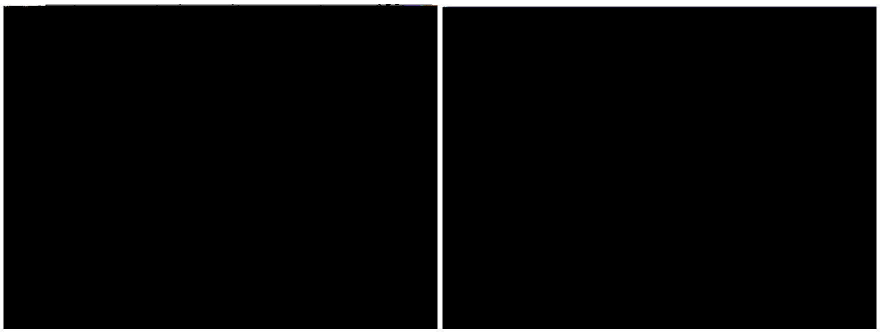

Fig. 2 Annual mean potential temperature increments (C month ^{1}) cross-section along equator between 140 E-90 W: (a) with standard data assimilation and (b) with the pressure correction scheme.

assimilation techniques, such as those described here, enable better use of these expensively acquired data to give more accurate predictions and reanalyses.

Accurate seasonal and decadal forecasts are particularly important for understanding the effects of climate change and in developing strategies for living with changes in our environment as well as for mitigating hazardous conditions that may arise, such as flooding, drought, intense rainfall, heavy snow and ice, or excessive temperatures. Increased accuracy of ocean, weather and climate forecasting has impacts on economic, commercial and organisational elements of society as well as on our understanding of the environment.

The development of the pressure correction method by the University of Reading and the Met Office, and the improvement in accuracy it has allowed, has made an important contribution to ocean data assimilation. The technique forms an integral part of the Met Office ocean data assimilation system. It has also been extended by ECMWF and incorporated into their ocean-atmosphere assimilation system for seasonal forecasting and plays an important role in their system for re-analysis of ocean heat uptake [1], a major factor affecting climate change.

Both the Met Office and ECMWF have now transitioned to a new coreocbcore

References

1. Balmaseda, M.A., Trenberth, K.E. and Kallen, E.:Distinctive climate signals in reanalyses of global ocean heat content. Geophys. Res. Lett., **40**, 1754–1759 (2013).
2. Bell, M.J., Martin, M.J., Nichols, N.K.: Assimilation of data into an ocean model with systematic errors near the equator. Quart. J. Roy. Meteorol. Soc., **130**, 873–894 (2004)
3. Gill, A.E.: Atmosphere-Ocean Dynamics, Academic Press, San Diego, London (1982)
4.