Available online at www.sciencedirect.com
ScienceDirect
Ocean Modelling 20 (2008) 1–16
Ocean Modelling
www.elsevier.com/locate/ocemod

Combining Argo profiles with a general circulation model in the North Atlantic. Part 1: Estimation of hydrographic and circulation anomalies from synthetic profiles, over a year

Gaël Forget [a,*], Bruno Ferron [b], Herlé Mercier [b]

[a] *Room 54–1517, Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b] *Laboratoire de Physique des Océans, UMR 6523, CNRS-Ifremer-UBO, Ifremer centre de Brest B.P. 70, 29280 Plouzané, France*

Received 12 January 2007; received in revised form 1 June 2007; accepted 4 June 2007
Available online 28 June 2007

Abstract

Argo is a global array of profiling floats that provides temperature (*T*) and salinity (*S*) profiles from 2000 m to the surface every ten days with a nominal spatial resolution of 3°. Here we present idealized experiments where the adjoint method is used to synthesize simulated sets of Argo profiles with a general circulation model, over a one-year period, in the North Atlantic. Using a number of drifting profilers consistent with Argo deployment objectives, the simulated array permits one to identify large-scale anomalies in the hydrography and circulation, despite the presence of a simulated eddy noise of large amplitude. Model dynamics provide an objective means to distinguish eddy noise from large-scale oceanic variability, and to infer the absolute velocity field (including abyssal velocities and sea surface height) from sets of Argo profiles of *T* and *S*. In particular, our idealized experiments suggest that volume and heat transports can be efficiently constrained by sets of Argo profiles. Increasing the number of Argo floats seems to be an adequate strategy to further reduce errors in circulation estimates.

Keywords: Data assimilation; 4DVAR; Twin experiment; Argo; General circulation model

1. Introduction

Argo is a global array of profiling floats, which provides temperature (*T*) and salinity (*S*) profiles from 2000 m to the surface every ten days, with a nominal spatial resolution of 3°. This subsurface observing system was designed to complement satellite observations (e.g. of sea surface height) in order to monitor the large-scale and low-frequency oceanic variability (see Roemmich et al., 1999).

* Corresponding author. Tel.: +1 617 253 5458; fax: +1 617 253 4464.
E-mail address: gforget@mit.edu (G. Forget).

1463-5003/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ocemod.2007.06.001

One way to investigate the adequacy of Argo as a large-scale observing system is to test whether it permits one to estimate large-scale T and S signals. Argo does not fully resolve the high frequency and small-scale oceanic variability (e.g. meso-scale eddies or internal waves), which thus might be aliased. This suggests the use of state estimation methods that are able to distinguish between large-scale signals of interest and small-scale *noise*. Guinehut et al., 2002, for example, show that the eddy noise can be efficiently filtered out to estimate monthly T and S anomalies in $6° \times 6°$ squares, using objective analysis (Bretherton et al., 1976). The adequacy of Argo as a large-scale observing system can also be investigated with respect to large-scale circulation fields besides T and S. As each pair of profiles can be interpreted as a direct measurement of the geostrophic shear, Argo can be expected to significantly improve our knowledge of the large-scale circulation through the use of inverse methods. Unlike previous studies, the adequacy of Argo is here investigated in terms of both the hydrography (i.e. T and S) and the circulation, with an emphasis on the latter.

The estimation method employed here, known as the "adjoint method" or the "4DVAR assimilation" (see e.g. Wunsch, 1996, and Bennett, 2002), consists of searching for the solution of a general circulation model (GCM) that best fits the observations in a least-squares sense, over a time interval. In this context, a GCM can be interpreted as a covariance operator, which accounts for a variety of length scales (from the model grid-scale, up to the model domain size), for a variety of time scales (from the model time-step, up to the simulated time interval), for the dynamical relationships between the different physical quantities, and for the anisotropy associated with oceanic structures. In particular, the GCM-based interpolation tool can be regarded as a non-linear inverse model in which the velocity field is estimated along with the tracer fields. The GCM-based interpolation framework is therefore appropriate to investigate the adequacy of Argo as a large-scale observing system as a whole, i.e. with respect to both the hydrography and the circulation.

We want to provide insights into the two following problems: by processing sets of Argo profiles with GCM-based interpolation, can the large-scale T and S signals be distinguished from the small-scale noise? To what extent can the large-scale circulation be inverted from sets of Argo (T and S) profiles?

The complexity of the problem argues for a two-step approach. Indeed, assume that only real-observation experiments are carried out, and the resulting Argo-GCM estimates (X_e, for a quantity X) are evaluated based on independent information (X_i). X_e could be a poor estimate of X_i if the information content of sets of Argo profiles was insufficient: (issue A) X may be dominated by scales that are too small compared with the Argo sampling; (issue B) if X is velocity information, it simply may not be possible to determine X uniquely from sets of T and S profiles. But, in practice, X_e could also be a poor estimate of X_i if the estimation system had faults: (issue C) the least-squares solution (X_e) may be incorrect because of improper assumptions on error statistics; (issue D) errors in the GCM dynamics may prevent the adequate fit to Argo profiles. If X_e was a poor estimate of X_i, it would then be unlikely that a clear explanation emerges. Now, idealized experiments using GCM simulated observations permit one to investigate issues A and B alone, leaving issues C and D as the main concerns when assimilating real observations. A two-step approach is therefore used: first an idealized context is considered where data noise statistics and GCM dynamics are assumed perfect (here), and then the case of real-observations is treated (companion paper: Forget et al., 2007).

Also, idealized experiments provide unique opportunities to investigate issues A and B. Indeed, the time-varying GCM state to be estimated (formally: *the truth*) is known completely and perfectly in idealized experiments, in contrast to the case of real oceans when X_i can be fairly uncertain, incomplete or simply non-existent. For example, idealized experiments allow one to test the possibility of inferring anomalies of volume and tracer transports from Argo measurements (Section 3). Furthermore, the data set characteristics are fully adjustable parameters in the idealized framework, and this permits a discussion of observing system design. For example, one can test whether increasing the number of Argo floats could significantly improve circulation estimates (Section 4).

2. Experiment configurations

2.1. Estimation problem

Consider an estimation problem that consists of finding a set of model parameters, x, that yields a least-squares fit of the model trajectory to noisy data, y_0, and to an uncertain prior estimate of x, x_b. The model