



# On the use of current meter data to assess the realism of ocean model simulations

Thierry Penduff, Bernard Barnier, Jean-Marc Molines, Gurvan Madec

## ► To cite this version:

Thierry Penduff, Bernard Barnier, Jean-Marc Molines, Gurvan Madec. On the use of current meter data to assess the realism of ocean model simulations. *Ocean Modelling*, 2006, 11 (3-4), pp.399-416. 10.1016/j.ocemod.2005.02.001 . hal-00124812

**HAL Id: hal-00124812**

**<https://hal.science/hal-00124812>**

Submitted on 11 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# On the use of current meter data to assess the realism of ocean model simulations

Thierry Penduff <sup>a,\*</sup>, Bernard Barnier <sup>a</sup>, Jean-Marc Molines <sup>a</sup>, Gurvan Madec <sup>b</sup>

<sup>a</sup> *Laboratoire des Ecoulements Géophysiques et Industriels, (CNRS-UJF-INPG), BP53, 38041 Grenoble Cedex 9, France*

<sup>b</sup> *Laboratoire d'Océanographie et de Climatologie: Expérimentation et Analyse Numérique  
(UPMC-CNRS-IRD), Institut Pierre Simon Laplace Tour 14-15 – 2ème étage – Boîte 100 – 4, place Jussieu,  
75252 Paris Cedex 05, France*

The evaluation of ocean simulations against observed datasets is essential to assess their realism and to guide model development, but often remains qualitative, and ignores certain datasets. This paper presents a three-dimensional, quantitative comparison of a  $1/6^\circ$  Atlantic numerical simulation (CLIPPER) with the WOCE current meter dataset in terms of mean velocity and eddy kinetic energy. Our metrics reveal the good behaviour of CLIPPER open boundary conditions and forcing with respect to full-depth current records. Due to its still moderate resolution, however, the model globally underestimates the observed mean speeds and eddy activity. This discrepancy is barely noticeable at low latitudes but increases toward the poles, probably since the poleward decrease of the Rossby radius exceeds that of the horizontal grid step. At least in this eddy-admitting regime, it is suggested that the numerics of geopotential-coordinate models like ours dissipate mean and eddy momentum at depth and adversely affect current–topography interactions.

---

\* Corresponding author. Tel.: +33 47682 50 65.  
E-mail address: thierry.penduff@hmg.inpg.fr (T. Penduff).

## 1. Introduction

So-called “realistic” ocean numerical models solve the primitive equations within realistic basin geometries, initial, lateral and surface boundary conditions, but the *actual* realism of such numerical solutions must be assessed a posteriori by comparing them to complementary datasets. Simulations are performed at increasingly finer resolutions over increasingly long periods, so that their validation requires the development of quantitative and synthetic tools. The definition and significance of model-data misfits need to be adapted to the nature and limitations of modelled and observed datasets. Model-data comparisons are generally based on means and variances derived from observed and modelled datasets rather than raw datasets themselves, since the events simulated without data assimilation are not necessarily in phase with observations (especially at mesoscale). Model validation studies often remain qualitative or partly quantitative. An exacting way to quantify the realism of a numerical solution is likely to diagnose and compare the same quantities from observed and simulated datasets at the same locations and instants (e.g. McClean et al., 1997, 2002; Tokmakian and McClean, 2003; Stammer et al., 1996). A multi-year, sustained, near-global monitoring of sea-surface height and derived anomalous velocity field is being performed by altimeters at “eddy-permitting” temporal and spatial resolutions. These data are essential and widely used to validate modelled dynamical fluctuations near the surface.

Datasets are much sparser below the surface, and basin-scale model-derived circulation schemes may be compared there to Eulerian maps based on binned subsurface drifter trajectories along surfaces of constant density or pressure (e.g. Tréguier et al., in press). The best available description of absolute mean and eddy flows throughout the water column is provided by current meter (CM) datasets. These datasets are local and sparse since relatively few instruments sample the three-dimensional ocean at fixed locations over limited time periods. To fully extract the dynamical information available from CM datasets requires dedicated, regional, thorough investigations (e.g. Arhan et al., 1989; Colin de Verdière et al., 1989; Woodgate et al., 1999) which clearly lie beyond the scope of the present study. CM velocity datasets may be of great interest also to validate full-depth model solutions.

The current meter data collected during the World Ocean Circulation Experiment (WOCE) are used in the present paper to estimate the main dynamical biases of a numerical simulation of the Atlantic Ocean at  $1/6^\circ$  resolution driven over the period 1979–2000 by reanalyzed atmospheric data (CLIPPER project, Tréguier et al., 1999). These model results have already been compared to observations and used for various dynamical studies in several papers (e.g. Tréguier et al., 2002; Candela et al., 2003; Hall et al., 2004; Penduff et al., 2004). The present paper proposes a method to quantify the differences between simulated and observed CM datasets in terms of mean velocity field and associated variances (eddy kinetic energy, EKE) over an ocean basin. The significance of these misfits is evaluated, and physical interpretations of the model strengths and weaknesses revealed by this novel type of model-data comparison are proposed. In Section 2, we describe the numerical model and CLIPPER  $1/6^\circ$  configuration. The treatment applied to the observed and simulated velocity fields and the subsequent statistics are described in Section 3. Comparison procedures and results are presented in Section 4. Model skills, computed as defined by Holloway and Sou (1996), are computed, and related with our own metrics in Section 5. A summary is given in Section 6.

## 2. Model configuration

CLIPPER is a French contribution to WOCE (World Ocean Circulation Experiment) consisting in the modelling of the Atlantic circulation driven by air–sea fluxes during the WOCE era (1979–2000). The CLIPPER project team has implemented several numerical configurations at various resolutions, based on the same primitive equation, rigid lid, geopotential-coordinate model (OPA8.1, Madec et al., 1998). The  $1^\circ$  and  $1/3^\circ$  Atlantic configurations are described in Beismann and Barnier (2004), and Tréguier et al. (2001), respectively. In the present study, we make use of the  $1/6^\circ$  simulation labelled “HF” in Penduff et al. (2004), whose features are briefly summarized hereafter. As shown in Fig. 1, the numerical domain is limited by radiating/relaxing open boundary conditions at the Drake Passage ( $68^\circ\text{W}$ ), south of South Africa ( $30^\circ\text{E}$ ), at the Gulf of Cadiz ( $8^\circ\text{W}$ ), and along  $70^\circ\text{N}$ . The performance of these boundary conditions was proven satisfactory by Tréguier et al. (2001). The bottom topography is based on the Smith and Sandwell (1997) database. The resolution of the isotropic horizontal grid is  $\Delta = 1/6^\circ \cos(\text{latitude})$ . Biharmonic horizontal viscosity and diffusion operators are used with coefficients  $A_h$  proportional to  $(\Delta/\Delta_{\max})^3$  ( $\Delta = \Delta_{\max}$  and  $A_h = 5.5 \times 10^{10} \text{ m}^4 \text{ s}^{-1}$  at the equator). Forty-two grid levels are used in the vertical, with a vertical resolution decreasing from 12 m at the surface to 200 m below

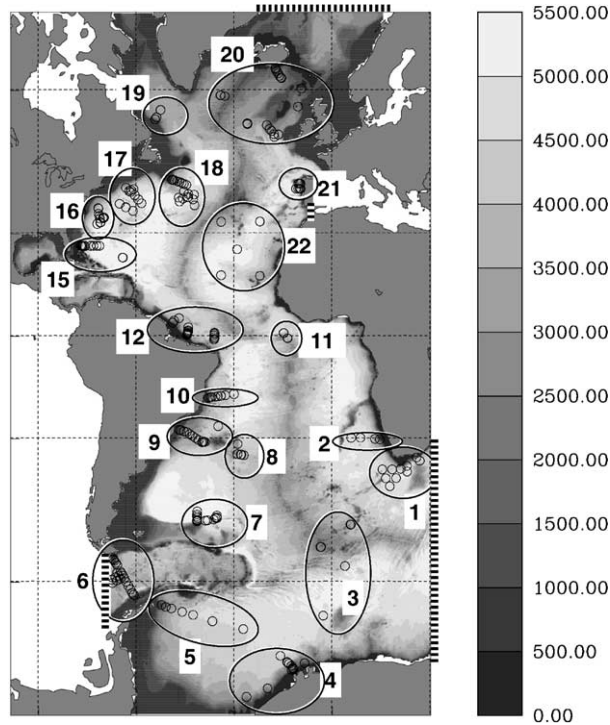


Fig. 1. Black marks show the location of WOCE moorings and of their CLIPPER model counterparts. Indexed ellipses identify the regional clusters used in this study (indexes 13 and 14 are not attributed). Model topography (m) is shown in the background, and the four open boundaries are shown as thick dashed lines south of America, Africa, in the Gulf of Cadiz and along the northern limit.

1500 m. Vertical viscosity and diffusion coefficients are given by a second-order closure scheme (Blanke and Delecluse, 1993), and are enhanced in case of static instability. Aliasing of high-frequency signals is avoided by storing the model outputs as successive 5-day averages (Crosnier et al., 2001). Surface forcing is applied as described in Barnier (1998). The model is started from rest, initialized by Reynaud et al. (1998)’s temperature and salinity seasonal climatology. It is spun-up for 8 years with a climatological seasonal forcing derived from the 1979–1993 ERA15 ECMWF reanalysis. The model is then forced successively by 1979–1993 ECMWF reanalyzed fluxes and 1994–2000 ECMWF analyzed fluxes (both interpolated at every timestep and consistently linked to each other, see Penduff et al., 2004). The reader is referred to this latter paper for more details about this numerical simulation.

### 3. Processing of observed and simulated datasets

#### 3.1. Processing of WOCE current meter data

In this study we make use of 1300 current meter measurements collected between 1979 and 2000 in the Atlantic (Fig. 1) during the World Ocean Circulation Experiment (WOCE). Sampling frequencies and record lengths range between 15 min to 12 h and from a few months to about 840 days, respectively. In order to build a homogeneous dataset comparable to the CLIPPER model outputs, these raw velocity measurements were selected, filtered, and formatted as follows. WOCE ( $u, v$ ) raw time series were averaged over successive 5-day intervals to build low-pass filtered ( $U, V$ ) time series. Error flags were inserted in the filtered time series if more than 30% of raw data were erroneous within 5-day windows. Resultant isolated flags were replaced by linear interpolation. Longer (2 or more) sequences of flags were not replaced. Continuous sequences of ( $U, V$ ) were then separated and linearly detrended to avoid aliasing of low-frequency velocity fluctuations. Resulting ( $U, V$ ) time series shorter than 6 months were excluded. Particular cases were checked individually, and the final dataset was globally verified.

This process retained 69% of the original WOCE data, thus provided 891 low-passed filtered, linearly detrended, continuous velocity time series longer than 6 months. As many estimates of averaged velocity components ( $U, V$ ), speeds  $|U|$  and EKEs were then deduced from WOCE time series at each location and over each recording period  $[t_1, t_N]$ . EKEs were computed as

$$\text{EKE}_{[t_1, t_N]} = \frac{1}{N} * \sum_{t=t_1}^{t_N} \frac{(U(t) - \overline{U}_{[t_1, t_N]})^2}{2} + \sum_{t=t_1}^{t_N} \frac{(V(t) - \overline{V}_{[t_1, t_N]})^2}{2},$$

where the overbar denotes the temporal mean over the available intervals  $[t_1, t_N]$ . The subsequent WOCE  $[\overline{U}, \overline{V}, \overline{|U|}, \text{EKE}]$  data are largely dispersed in space and time. Fig. 2a shows that data are available throughout the water column over the whole model integration but irregularly distributed in time (i.e. data are more abundant over the period 1990–1995 than during 1981, 1988, 1989, and after 1996). The median length of selected WOCE records does not appear to depend much on depth (Fig. 2b), but significantly increases from the 1980s to the 1990s (close to 1 and 1.3 years respectively, not shown). Our model-data comparison will be performed by depth range, and within 20 geographical clusters (red circles in Fig. 1).

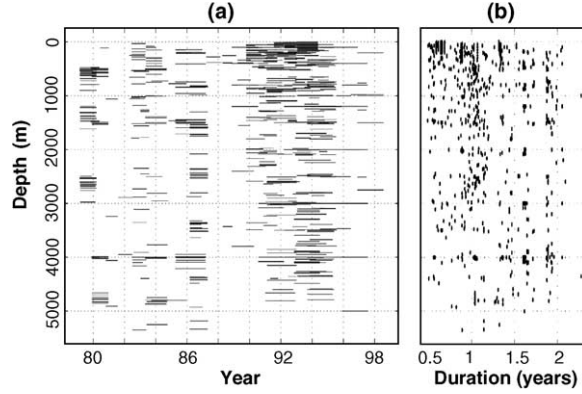


Fig. 2. (a) Vertical and temporal distribution of the 806 available synoptic, co-located velocity time series extracted from the WOCE and CLIPPER current meter datasets throughout the Atlantic. Lines show the immersion, starting date and duration of the time series. (b) Depth of every time series (m) as a function of their duration. Minimum duration is set to 6 months (0.5 year).

### 3.2. Processing of CLIPPER model outputs

CLIPPER outputs were saved as successive averages of model fields over 5-day periods between 1979 and 2000. Model counterparts of every WOCE ( $U$ ,  $V$ ) individual time series were extracted and detrended from CLIPPER outputs at the closest model point and time period. Model outputs were not interpolated at the exact current meter locations to match the very local character of WOCE measurements. The distance between real and simulated current meters remains small: less than  $1/12^\circ \times \cos(\text{latitude})$  in the horizontal and 100 m in the vertical. The extraction and detrending of model velocities was not done at 85 coastal or near-bottom sites where the closest model grid points are masked by the discrete topography. Model mean velocity components  $[\overline{U}, \overline{V}]$ , speeds  $|\overline{U}|$  and EKE estimates were then computed as WOCE data to finally provide 806 quasi-located synoptic pairs of simulated and observed  $[\overline{U}, \overline{V}, |\overline{U}|, \text{EKE}]$  estimates. Model and observed statistics may be compared consistently since they derive from quasi-located ( $U$ ,  $V$ ) time series with the same low-pass filtering, temporal sampling, and local character.

Observed and simulated  $[\overline{U}, \overline{V}, |\overline{U}|, \text{EKE}]$  estimates are representative of a limited time period and cannot be assumed constant over longer timescales. For example, a local 0.25 m/s difference between observed and simulated “mean” current speeds will be particularly significant in regions where current speeds barely vary on interannual timescales. The variability of  $[\overline{U}, \overline{V}, |\overline{U}|, \text{EKE}]$  over timescales longer than recording durations should thus be quantified to properly define model-data misfits. This information is unknown in the real ocean and was estimated from the model outputs as follows. At every model mooring (Fig. 1) and vertical level  $k$ , the 1980–2000 detrended model velocity time series were split into 10 successive 2-year segments to compute as many 2-year simulated estimates of  $W$  (where  $W$  designates any variable in  $[\overline{U}, \overline{V}, |\overline{U}|, \text{EKE}]$ ). The distributions of these estimates were computed at every cluster and model level  $k$  to provide depth-dependant, regional estimates of the median (noted  $\langle W \rangle$  for each variable  $W$ ), of the 17th and 83rd percentiles (noted  $W_{17\%}$  and  $W_{83\%}$  respectively) of each distribution. Intervals of

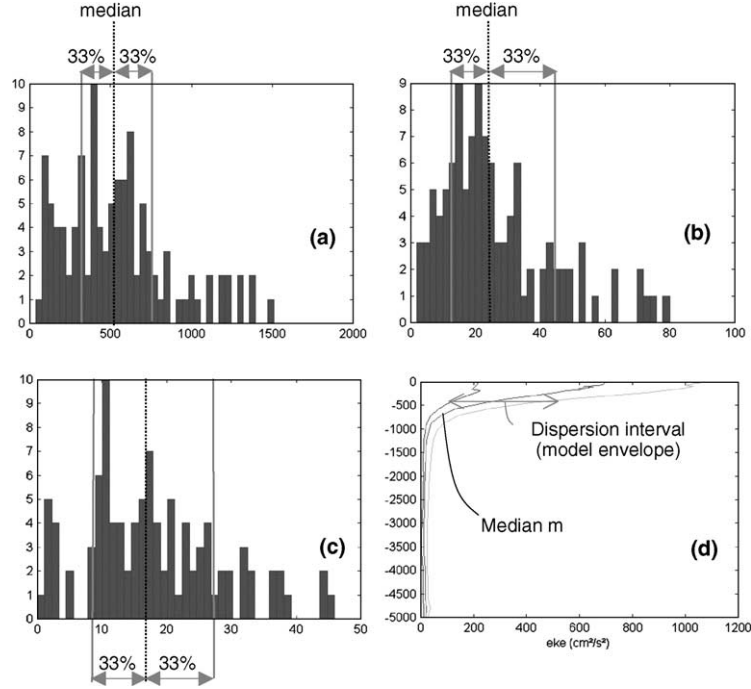


Fig. 3. Model EKE distributions at different immersions ((a) 227 m; (b) 1417 m; (c) 4400 m) computed from the 1980–2000 time series split into 10 segments at cluster 17. This cluster is located in the Gulf Stream. Abscissas: EKE in  $\text{cm}^2/\text{s}^2$ ; ordinates: number of values. Dashed vertical lines show the median EKE, plain lines show the lower and upper limits of the 17–83% quantile, i.e. the centered dispersion intervals which include 66% of values. Panel (d) presents the full-depth structure of the EKE median and dispersion interval in this particular cluster.

time-space dispersion around model medians are defined as  $|W_{83\%} - W_{17\%}|$ , and will be also referred to as “model envelopes” in the following. Fig. 3 shows examples of model distributions, medians and envelopes (EKE at cluster 17) at selected depths (panels a, b, and c) and their resulting vertical structure (panel d). The dispersion within clusters was generally found to be more temporal than spatial, confirming the dynamical homogeneity of these clusters.

Model-derived vertical profiles of  $[\overline{U}, \overline{V}, \overline{|U|}, \text{EKE}]$  medians, percentiles and envelopes are shown for in Fig. 4 at three typical clusters, along with paired WOCE and CLIPPER individual estimates (marks). As expected, 66% of individual CLIPPER values (+ and × in Fig. 4) fall within model envelopes at the same cluster. Fig. 4 illustrates several features of the datasets.

### 3.3. Description of the datasets

The quantity and vertical distribution of WOCE measurements differ from one cluster to another. Data are dense and relatively uniform in the vertical in the western equatorial region (cluster #12), but they are sparser and more irregularly distributed along the Falkland Escarpment (#7) and the Antarctic shelf (#4), i.e. absent over hundreds of meters in the eddy-active surface layer and at depth, respectively. This inhomogeneity should be taken into account for the interpretation of model-data differences.

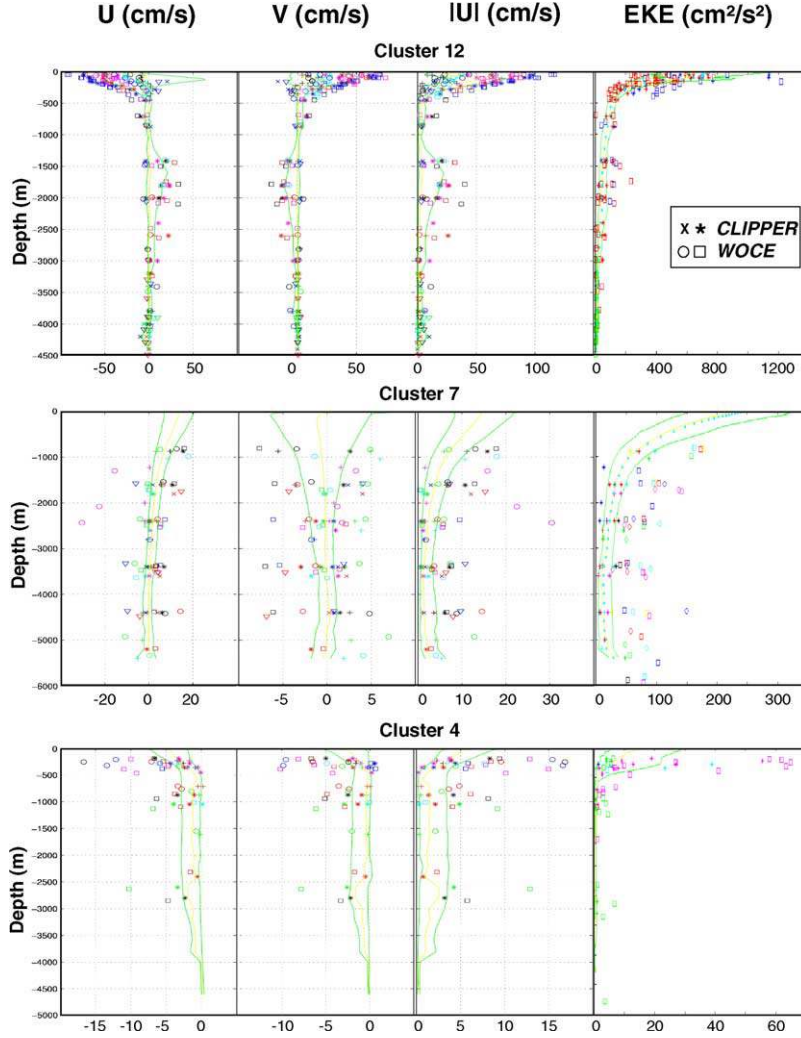


Fig. 4. Vertical structure of WOCE (circles, squares) and CLIPPER (crosses, pluses) estimates of  $U$ ,  $V$ ,  $|U|$  and EKE (from left to right column, respectively) in clusters 12, 7 and 4 (upper, middle, and lower panels, respectively). Marks with identical color correspond to synoptic, co-located estimates. Superimposed model medians and envelopes are deduced from the 21-year simulation are shown as yellow and green lines.

The lines in Fig. 4 show a complex, depth-dependant and regional diversity in  $[\overline{U}, \overline{V}, \overline{|U|}, \text{EKE}]$  model medians and envelopes. For example, median  $\langle \overline{U}, \overline{V} \rangle$  profiles show that time-averaged vertical shears are stronger at low latitudes than at higher latitudes, as expected from the equatorward increase of stratification.

Model  $[\overline{U}, \overline{V}, \overline{|U|}, \text{EKE}]$  individual estimates visibly resemble or differ from their WOCE counterparts in many ways: the model-data misfit should be estimated from complementary criteria. For example, one may distinguish between the model biases visible throughout the water column (such as the ones visible on  $[\overline{|U|}, \text{EKE}]$  at cluster #7 or on  $[\overline{U}, \overline{V}]$  at cluster #4) and those limited to certain depth ranges.



#### 4. Comparison procedures and results

The model realism in terms of  $[\overline{U}, \overline{V}, |\overline{U}|, \text{EKE}]$  is evaluated by cluster and depth range in the present and following sections. After a qualitative description of the observed and simulated current meter datasets in Section 4.1, individual model-data misfits and model envelopes are used in Section 4.2 to define quantitative “model agreement” indexes, taking into account the four-dimensional sparsity and dispersion of the datasets.

##### 4.1. Speed, EKE, baroclinicity

The observed and simulated datasets were first split in two subsets, comprising data above and below 1000 m (both resulting subsets have similar sizes). The relative difference between upper and lower median speeds  $|U|$  and EKEs will be referred to as “baroclinicity” hereafter. Fig. 5 shows for each cluster and both depth ranges the WOCE and CLIPPER median speeds  $|U|$  and EKEs.

CLIPPER speeds and EKE levels are globally comparable with, but generally weaker than their WOCE counterparts. The modest model resolution, the classical use of biharmonic horizontal

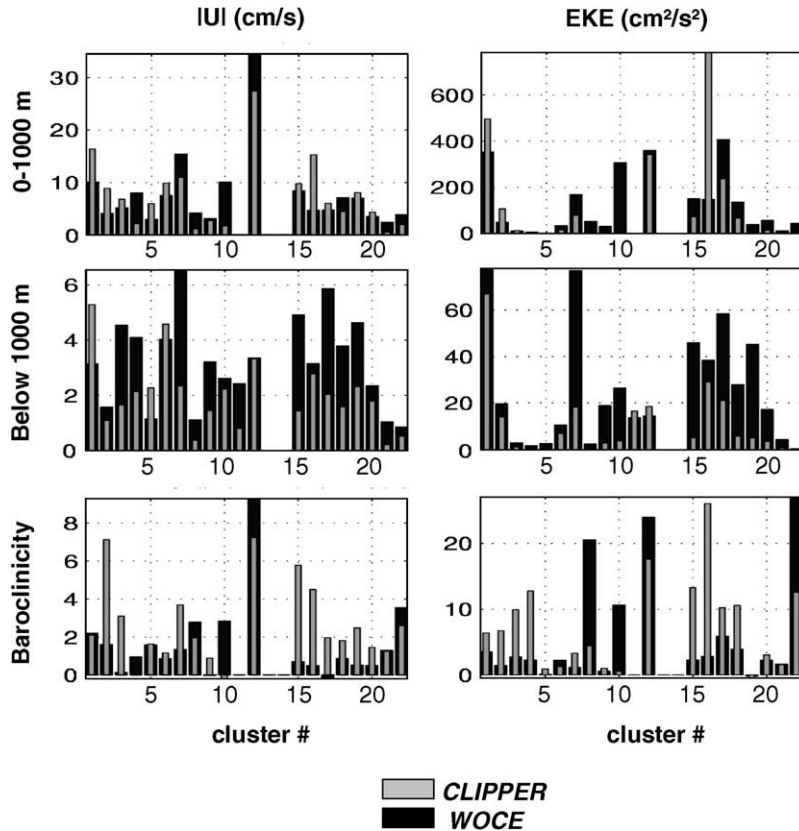


Fig. 5. Median speeds  $|U|$  (left) and EKEs (right) above 1000 m (upper panels) and below 1000 m (middle panels) at each cluster (cluster index indicated in abscissas). Lower panels show the baroclinicity of both quantities at each cluster, i.e.  $[\text{above } 1000 \text{ m} - \text{below } 1000 \text{ m}]/[\text{below } 1000 \text{ m}]$ . CLIPPER results are shown in gray and WOCE results in black.

friction and of forcing fields at relatively coarse temporal and spatial resolutions (daily,  $1.125^\circ$ ) are all expected to keep the simulated dynamics in an eddy-admitting, too viscous regime. Despite exceptions at certain clusters (commented below), model speeds  $|U|$  are more comparable with WOCE estimates above 1000 m than below, where they are clearly underestimated. This leads to an overestimated baroclinicity of  $|U|$  in CLIPPER (lower left panel in Fig. 5) within most clusters, especially at high latitudes in the North Atlantic (clusters 15–20). The simulated EKE field also exhibits a baroclinic discrepancy (lower right panel). Overestimated baroclinicities of  $|U|$  and EKE may be explained in two ways.

1. The bias in  $EKE(z)$  was previously discussed by Penduff et al. (2001, 2002, noted P1-2 in the following) in a regional configuration of the same model at coarser resolution ( $1/3^\circ$ ). P1-2 suggested that numerical discrepancies inherent to geopotential-coordinate models like OPA may adversely dissipate momentum along topographies and thus exaggerate  $|U|$  and EKE baroclinicities. If that is true, a better representation of mesoscale current–topography interactions (due either to increased resolution or better numerical formulation of topographic constraints) should moderate the bias on both variables at the same time. This is true in the present simulation: most exceptions to the general baroclinic bias are found simultaneously on EKE and  $|U|$  fields (clusters 8, 10, 12, and 22). In addition, these exceptions are found at relatively low latitudes (within  $30^\circ\text{S}$ – $30^\circ\text{N}$ ) where the model grid better resolves the Rossby radius (thus mean currents, non-linearities, turbulence, topography, current–topography interactions), and where strong stratification moderates topographic effects. The present results thus support P1-2’s hypothesis of an intrinsic model discrepancy affecting current–topography interactions, and suggest that horizontal resolution may limit its adverse consequences.
2. These baroclinic biases on  $|U|$  and EKE, along with the quasi-absence in the model solution of major topographically-locked eddy-driven circulation features (like the barotropic Zapiola anticyclone in the Argentine basin) might also be due to the overall lack of eddy energy. Indeed, current–topography interactions were shown to generate bottom-intensified (possibly barotropic) rectified flows over topographic slopes in the stratified ocean (Merryfield and Holloway, 1999; de Miranda et al., 1999). The lack of bottom-intensified or barotropic mean and eddy momentum diagnosed in the present simulation at mid and high latitudes might thus be a combined effect of the insufficient resolution of eddy scales there and of using operators that dissipate not only enstrophy but also energy. These effects, along with the scheme- or resolution-related issue mentioned above, may adversely affect current–topography interactions, and limit the emergence of kinetic energy at depth.

Both explanations highlight that improved numerics (and/or parameterizations) are necessary to simulate proper current–topography interactions in geopotential-coordinate models, at least at eddy-admitting resolution. This quite robust numerical problem is currently being investigated.

Median  $[\overline{U}, \overline{V}, EKE]$  values computed above and below 1000 m at each cluster from WOCE and CLIPPER individual estimates are displayed in Fig. 6 as vectors and circles (whose radii show  $EKE^{1/2}$ ). It confirms the main model biases deduced from Fig. 5, i.e. the general underestimation of current speeds and EKEs, especially at depth, with additional information on current

### Medians over clusters

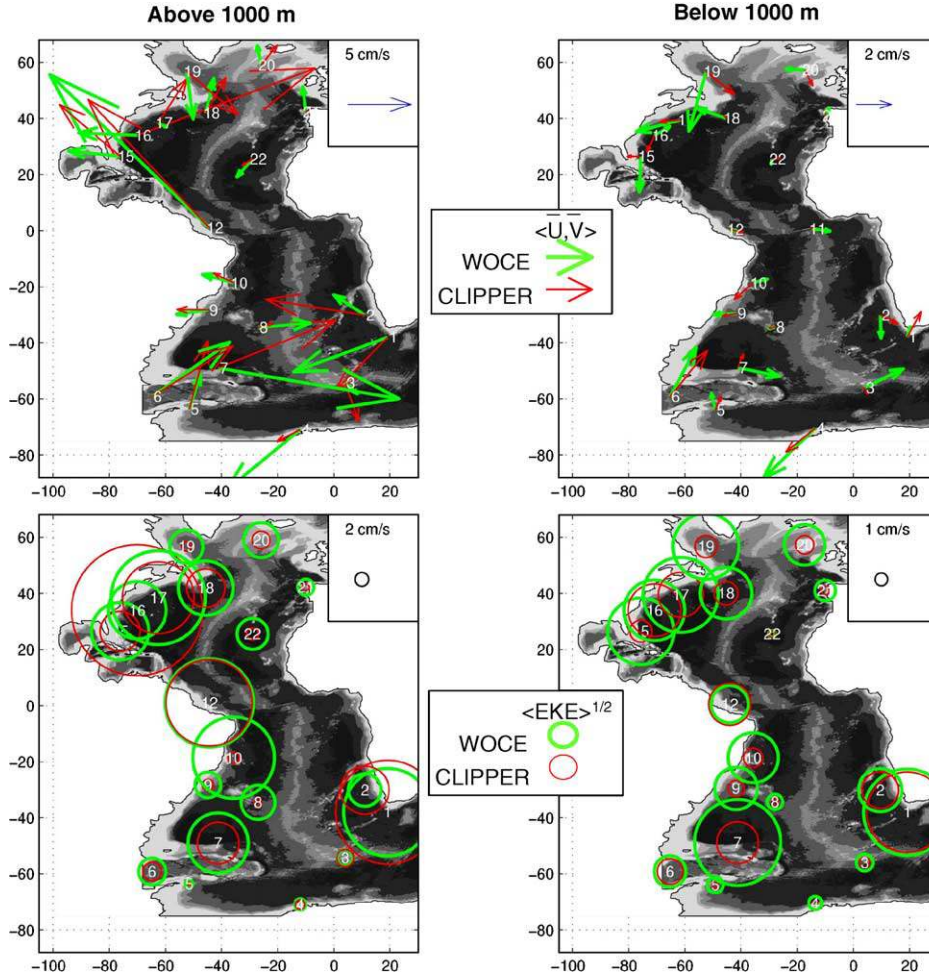


Fig. 6. WOCE (thick) and CLIPPER (thin) time-averaged velocities (upper panels) and speed standard deviations (lower) above (left) and below (right) 1000 m. Shown are median quantities computed over every cluster. Arrows start and circles are centered at the clusters' median locations.

directions. The model simulates particularly well the observed mean current direction, speed and EKE in the western equatorial cluster (#12). Despite rather high latitudes, very good agreement is also found near the Drake Passage (#6) and south of Africa (#1, #2). This confirms the good performance of CLIPPER open boundaries shown by Tréguier et al. (2001). The biggest discrepancy appears off Cape Hatteras (#16), with simulated currents heading opposite to observations in both layers and upper-layer median EKE five times as measured at the same cluster (Figs. 5 and 6). Indeed, the model Gulf Stream overshoots to the north, creating a strong, unrealistic, anti-cyclonic standing eddy, which is usual in geopotential-coordinate models at this resolution. This contaminates our results in cluster 16, which largely sits within the anticyclone's westward flow.

## 4.2. Quantitative estimates of model-data agreement

### 4.2.1. Definitions

*Absolute misfits* were computed as the differences between the 806 WOCE mean velocity and speeds  $(\overline{U}^W, \overline{V}^W, |\overline{U}|^W)$ , and the CLIPPER median values at the corresponding cluster  $p$  and nearest model level  $k$   $(\langle \overline{U}^C \rangle, \langle \overline{V}^C \rangle, \langle |\overline{U}|^C \rangle)$ . Absolute misfits are noted  $amu$ ,  $amv$ , and  $amU$  for  $\overline{U}$ ,  $\overline{V}$  and  $|\overline{U}|$ , respectively. For example, each meridional absolute misfit computed at location  $(p, k)$  reads as

$$amv_{p,k} = \overline{V}_{p,k}^W - \langle \overline{V}^C \rangle_{p,k}.$$

The width of model envelopes, derived within each cluster from the 17th and 83rd percentiles of CLIPPER statistics, provides regional depth-dependant estimates of the time-space dispersion of  $\overline{U}$ ,  $\overline{V}$ , and  $|\overline{U}|$  (see Section 3.2). Absolute misfits are then combined with model envelopes and medians to derive normalized random variables with non-normal statistics (noted  $rmu$ ,  $rmv$ , and  $rmU$  for  $\overline{U}$ ,  $\overline{V}$ , and  $|\overline{U}|$  respectively), that we called *relative misfits*. For instance, the meridional relative misfits at cluster  $p$  and vertical level  $k$  are computed as

$$rmv_{p,k} = \frac{amv_{p,k}}{\left| V_{p,k}^{83\%} - \langle \overline{V}^C \rangle_{p,k} \right|} \quad \text{where } amv_{p,k} > 0,$$

and

$$rmv_{p,k} = \frac{amv_{p,k}}{\left| \overline{V}_{p,k}^{17\%} - \langle \overline{V}^C \rangle_{p,k} \right|} \quad \text{where } amv_{p,k} < 0$$

(similar expressions hold for  $rmu$  and  $rmU$ ). Consequently, one gets  $-1 < rmv < 1$  for a WOCE  $\overline{V}$  estimate that falls within the CLIPPER envelope at the corresponding cluster and depth. One finds  $rmv = 0$ ,  $-1$ , and  $+1$  for an individual WOCE estimate of  $\overline{V}$  falling on the model median, on the lower, and on the upper limits of the dispersion interval, respectively. A simulated velocity biased to the south yields  $rmv > 0$ . The global and depth-dependant distributions of  $[\overline{U}, \overline{V}, |\overline{U}|]$  absolute/relative misfits are shown in Fig. 7. Table 1 gives the median misfits deduced from the global dataset (upper panels in Fig. 7).

Quantitative *model agreement* indexes were finally computed as the percentage of  $rmu$ ,  $rmv$ , and  $rmU$  estimates falling between  $-1$  and  $1$ : they correspond to the percentage of WOCE  $\overline{U}$ ,  $\overline{V}$ , and  $|\overline{U}|$  individual estimates falling within quasi-collocated model envelopes. Model agreement indexes are provided globally (last line in Table 1), by cluster (Fig. 7k) and by depth range (Fig. 7l).

### 4.2.2. Interpretation

Fig. 7d and e and Table 1 clearly show that the absolute  $U$ ,  $V$  and  $|U|$  misfits are distributed around very small median values. More precisely, roughly 40–45% of WOCE  $[\overline{U}, \overline{V}]$  and 50% of  $|\overline{U}|$  estimates fall within quasi-collocated model envelopes (Table 1). Fig. 7c and h and the global  $rmU$  median confirm the global underestimation of simulated current speeds at every

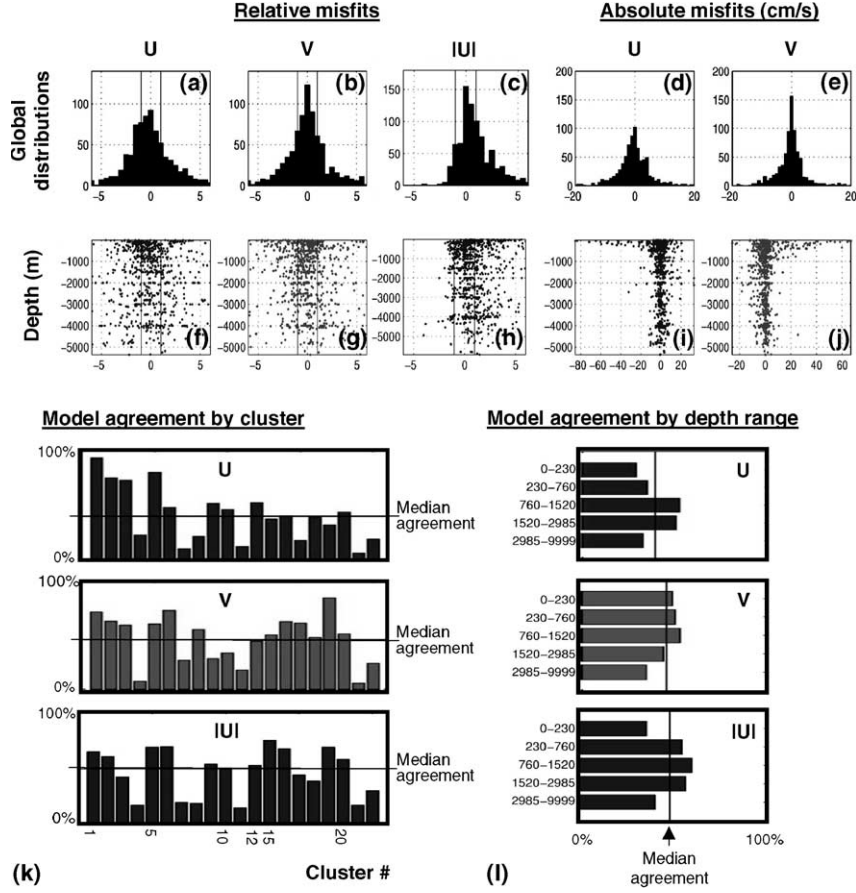


Fig. 7. Distribution of  $U$ ,  $V$ , and  $|U|$  relative misfits (upper left panels: a,b,c,f,g,h) and absolute misfits (upper right panels d,e,i,j). These distributions are shown globally and as a function of depth. Vertical lines in the upper left panel highlight the part of WOCE velocity estimates falling within collocated model envelopes, i.e. when relative misfits belong to  $[-1; 1]$ . Lower panels (k and l) show the model agreements for  $U$ ,  $V$ , and speeds  $|U|$  (from top to bottom), i.e. the percentage of WOCE estimates falling within corresponding model envelopes. These agreements are shown by cluster (k) and by depth range (l). Depth ranges in (l) were chosen so as to include a comparable number of available measurements. Plain lines in (k,l) show the “global model agreement” for each quantity (Table 1). See text for details.

Table 1

Medians (first line) deduced from the distribution of the 806 individual WOCE-CLIPPER relative (first three columns) and absolute (last two columns) misfits

	Relative misfits			Absolute misfits (cm/s)	
	$U(rmu)$	$V(rmv)$	$ U (rmU)$	$U(amu)$	$V(amv)$
Global median	-0.28	-0.01	0.7	-0.55	-0.01
Global model agreement	39%	46%	49%		

Since the absolute misfits are less meaningful than relative misfits, only  $amu$  and  $amv$  are given for illustration. The last line gives the global model agreement, i.e. the percentage of  $rmu$ ,  $rmv$ , and  $rmU$  values falling between  $-1$  and  $1$ , i.e. the fraction of WOCE  $U$ ,  $V$ , and  $|U|$  individual estimates falling within quasi-collocated model envelopes (see text and Fig. 7).

depth. The exaggerated baroclinicity of model speeds, deduced earlier from Fig. 5, also shows up in Fig. 7h since  $rmU$  values (i.e. the model speed underestimation) increase with depth.

Fig. 7k summarizes for each cluster the model agreement indexes for  $[\overline{U}, \overline{V}, |\overline{U}|]$  (noted  $Ax$ ,  $Ay$  and  $A$ , respectively), i.e. the percentage of WOCE estimates for each quantity falling within corresponding model envelopes. In most clusters,  $Ax$  and  $Ay$  are of similar magnitude, showing the isotropic nature of relative misfits. High (low)  $Ax$  or  $Ay$  values generally yield a high (low) speed agreement  $A$ , with the unexplained exception of cluster 3 in the Southern Ocean. With respect to their global medians (plain lines in Fig. 7k and percentages in Table 1),  $Ax$  and  $Ay$  model agreements are the best in the Agulhas Retroflexion (clusters 1, 2) and the ACC (3, 5, 6), showing again the good performance of southern open boundaries.

$Ax$ ,  $Ay$  and  $A$  happen to be particularly weak in clusters 4, 7, 11, 21 and 22. In the northern subtropical gyre (cluster 22) where the observations are confined in the upper 500 m, the weak agreement is explained by underestimated model speed, EKE and dispersion. Along the southern limit of the Weddell Sea (cluster 4) the simulated mean speeds and EKE are too weak as well, despite the realistic south-westward orientation and vertical structure of the current. Indeed, only few  $[\overline{U}, \overline{V}, |\overline{U}|, \text{EKE}]$  WOCE points fall within model envelopes there (cluster 4 in Fig. 4). This weak simulated circulation might be due to the climatological relaxation of tracers in this region.  $[\overline{U}, \overline{V}]$  WOCE estimates are satisfactorily centred on the model envelopes along the Falkland Escarpment (cluster 7, Fig. 4), the Iberian slope (#21) and the deep central equatorial basin (#11, not shown). Poor  $Ax$  and  $Ay$  agreement is also explained there by model speeds and EKE that are too weak and envelopes that are too narrow. The realism of the current directions found in clusters 7 and 21 might be favoured by the strong (realistic) topographic constrain exerted on the local flow by steep slopes.

The distribution of  $[\overline{U}, \overline{V}, |\overline{U}|]$  agreements was finally split into five subsets of similar size spanning the whole depth range (Fig. 7l). The percentage of WOCE  $|\overline{U}|$  estimates falling within model envelopes is maximum (about 60%) at intermediate depths and minimum (about 40%) within the bottom and surface layers. This suggests again the presence of spurious bottom friction at depth (poor  $A$  is explained by both  $Ax$  and  $Ay$ ), but also highlights a model bias in the upper 250 m, especially in the zonal direction. Indeed, in the top 250 m or so, some  $amu$  estimates reach strongly negative values (Fig. 7i) that are slightly anticorrelated (not shown) with associated  $amv$  values, suggesting south-eastward biases of upper model velocities at certain locations. When normalized as relative misfits, the southward component of this bias ( $rmv$ , Fig. 7g) decreases more than its eastward component ( $rmu$ , Fig. 7f), which is therefore more robust. Closer investigation reveals that this upper-layer discrepancy is confined downstream of Cape Hatteras (cluster 16, see Section 4.1 and Fig. 6). The slight underestimation of the powerful, north-westward North Brazil current near  $[44^\circ\text{W}, 0^\circ]$  (cluster 12, Fig. 6) also explains this eastward surface bias. This apparent global misfit is therefore mostly due to local discrepancies.

## 5. Model skills and synthesis

### 5.1. Definitions

“Skill” indexes were defined by Holloway and Sou (1996, noted HS96 hereafter) to estimate the realism of a model simulation against a sparse and spatiotemporally dispersed current meter

velocity dataset. In their study, local inner products of model and observed mean velocity vectors are weighted by the inverse of local observed EKEs, and summed up. HS96 then derive two indexes named *skillA* and *skillD*: the former quantifies the realism of simulated current vectors (directions and intensities) while the latter quantifies the realism of simulated current directions only. By construction, HS96's skill estimates equal 1 for a perfect model, 0 for a skill-less model (if model and observed  $[\bar{U}, \bar{V}]$ s were randomly unrelated), and would take negative values if model and observed vectors were opposed at numerous data points. HS96's directional and vectorial skill indexes complement our model speed agreement index *A*.

We computed *skillA* and *skillD* from the 806 CLIPPER/WOCE  $[\bar{U}, \bar{V}, \text{EKE}]$  individual estimates within five depth ranges (same as in Fig. 7l). As done by HS96 to assess the robustness of the results, *skillA* and *skillD* were computed 100 times within every subset, randomly rejecting (with probability 0.5) individual current meter records at each trial. The means and standard deviations of skills resulting from this procedure are shown respectively as thick lines and by the width of grey rectangles in Fig. 8. They hardly depend on small changes in the definition of depth ranges.

## 5.2. Interpretation and synthesis

As explained above, the modest speed agreement *A* obtained globally above 230 m (Fig. 7l) is due to local model discrepancies. Strong EKE there largely reduces the contribution of this bias in HS96's skills. The significant (narrow rectangles in Fig. 8) maxima reached above 230 m by *skillA* and *skillD* confirm the global realism of the model solution near the surface, and thus of the forcing fields, with correct orientations (*skillD* close to 0.5) and intensities (yet slightly underestimated as shown in Fig. 5). *skillA* decreases quasi-monotonically from top to bottom layers (Fig. 8a); our *A* index (Fig. 7l) and *skillD* (Fig. 8b) do not vary much at intermediate levels and reach slight local maxima at different depths (within the third and fourth layers respectively).

The poorest model performance is found in the topographically-influenced deepest layer (minimum of *skillA* below 2985 m): speeds are underestimated (as mentioned earlier, see Figs. 5 and 7l)

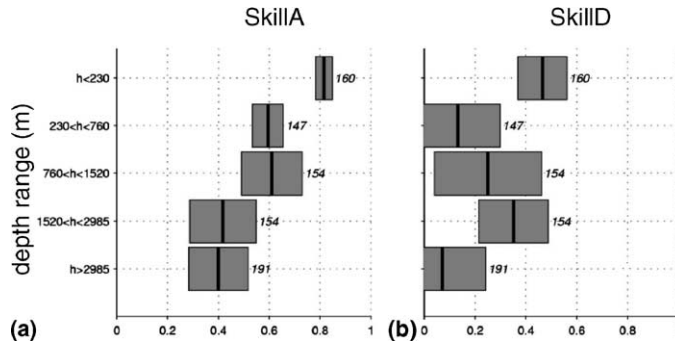


Fig. 8. Absolute model skill (*skillA*, panel a) and directional model skill (*skillD*, panel b). Results are computed as defined by Holloway and Sou (1996) and shown within the same depth ranges as in Fig. 7. Thick lines indicate the mean scores computed over 100 random discard trials performed over every subset independently. Standard deviations over these trials are indicated by the width of grey rectangles. Numbers next to rectangles indicate the number of WOCE/CLIPPER pairs used within each subset.

and directional skill is poor (Figs. 6 and 8b). This latter fact is confirmed by the weak circular correlation (defined as in Fisher and Lee, 1983) found within this depth range between modelled and observed velocity directions (about 0.02, i.e. 7–10 times less than in the uppermost three layers). The poor representation of topography as staircases is believed to adversely affect the intensities and directions of currents, especially at depth. Indeed, topographic influences increase toward the ocean bottom, especially at relatively small scales (as predicted by the Prandtl vertical scale). The present full-step topographic formulation and low vertical resolution at depth (up to 200 m) certainly distort the topographic details that actually steer the currents measured at deep WOCE locations. A good simulation of deep currents requires a better representation of topography (such as the “partial steps” discretization), in addition to the other suggestions made in this paper (improved numerics, higher resolution, if not dedicated parameterization).

The reader may have noted that we did not compute all performance indexes ( $A_x$ ,  $A_y$ ,  $A$ ,  $skillA$ ,  $skillD$ ) for every variable  $[\overline{U}$ ,  $\overline{V}$ ,  $|U|$ , EKE]. We did so to keep the paper short and synthetic, but for statistical reasons as well. Indeed, HS96’s skill indexes take into account the variance of the variables they are applied on (i.e. EKEs in the case of velocities). Unlike mesoscale velocity fluctuations (Gille and Smith, 2000), distributions of speeds  $|U|$  and EKEs are not close to symmetric. For this reason, the model performance in terms of speed  $|U|$  is quantified more consistently by our agreement index  $A$  which is better adapted to non-symmetric distributions. The same metrics may be applied to EKE as well, but was not evaluated in the present study. Also, we believe the model realism in term of velocity direction, which has a periodic distribution, is more properly evaluated by HS96’s  $skillD$  (based on inner products between observed and simulated normalized velocity vectors) than by any index comparable to our speed agreement  $A$ . Given the coarse and largely dispersed WOCE current meter dataset, and the unknown character of many features of the four-dimensional oceanic variability, the information provided by these complementary indexes need to be synthesised and interpreted carefully. However, the indexes computed in the present study lead to a rather clear, physically-consistent picture of the model’s behaviour, which complements more usual model validation exercises (see Tréguier et al., 2002; Candela et al., 2003; Hall et al., 2004; Penduff et al., 2004), and is summarized in the following section.

## 6. Conclusion

The quantitative validation of “realistic” numerical ocean simulations requires the reference to different types of complementary observed datasets and the development of adequate metrics. In this study, the WOCE current meter database and the 1/6°-resolution CLIPPER velocity fields were processed identically to provide 806 comparable pairs of synoptic, quasi-collocated estimates of mean velocity components, current speeds and EKEs, largely dispersed in time (1980–2000) and space (the three-dimensional Atlantic basin). The misfit between both datasets was quantified at each available location in terms of speed, velocity orientation and EKE. Depending on the distribution on these quantities, misfits were evaluated with originally-defined skill estimates (HS96), or with “model agreement” estimates based on the dispersion of simulated quantities around their median over 20 years. Model skill indexes were applied to validate mean velocity vectors and their direction, while agreement indexes were applied to quantities like current speeds whose distributions are asymmetric (EKE agreements may be evaluated this way as well). Skills and agreements



were computed as a function of depth or geographical location (clusters) to localize the model's strengths and weaknesses.

Model fields were found to agree well with WOCE data near the Agulhas retroflexion region and the Drake Passage, confirming the satisfactory behaviour of CLIPPER open boundary conditions. As generally mentioned at comparable or even higher resolution (Maltrud and McClean, 2005), our comparison reveals however the general underestimation of simulated speeds  $|U|$  and EKEs, and thus highlights the need for more consistent and/or less dissipative numerics (improved schemes, higher resolution, more selective eddy viscosity operator), and perhaps more energetic forcing fields (high wavenumber/frequency winds for instance). The contribution of an improved forcing could be quantified from the same metrics applied to sensitivity experiments. The realism of mean current directions was found very poor at depth, probably because of the inadequate full-step representation of topographic details and their steering influence. At mid and high latitudes, i.e. where the model is in the eddy-admitting regime, the underestimation of CLIPPER speeds and EKEs gets more severe with increasing depth. There, inconsistent numerics are suspected to generate spurious near-bottom friction and induce these baroclinic biases (as hypothesized by Penduff et al., 2001, 2002). On the other hand, our classical (biharmonic) viscosity operator that diffuses momentum down-gradient is expected to weaken eddy-topographic interactions (Merryfield and Holloway, 1999) and thus impede the generation of deep momentum. Depth-dependant profiles of simulated speeds and EKEs get increasingly realistic with decreasing latitude (as seen in the equatorial region). This improvement is consistent with the stronger local stratification which confines topographic influences and discrepancies at depth, and with the better resolution of low-latitude, i.e. much larger, internal deformation radii (thus of mesoscale topographic and dynamical features). This latter feature strongly suggests that better resolving the natural scales of motion improves the representation of topographic impacts. Numerical improvements thus appear necessary to properly simulate topographic influences at eddy-admitting resolution.

Present models essentially differ through their vertical coordinate systems and their formulation of topography, and their simulated mean and eddy flows can be radically different (Willebrand et al., 2001; Chassignet et al., 2000; Barnier et al., 2001; Penduff et al., 2001). Vertical profiles of oceanic properties reflect a number of important processes (forcing and topographic impacts, intermediate circulation, inverse cascade) and should thus be considered in model intercomparison exercises, as well as in single model validations. Current meter observations are necessary for that. Improved numerical schemes, better formulations of lateral/bottom boundary conditions, and/or additional subgrid-scale parameterizations are required for a better representation of current-topography interactions and of their impact on the water column at eddy-admitting resolution. This is especially the case in geopotential-coordinate models, where the use of partial or shaved cells (Adcroft et al., 1997; Pacanowski and Gnanadesikan, 1998) may solve part of the problem. It is likely that topography and bottom boundary conditions are more naturally formulated in sigma- than in geopotential-coordinate models (bottom friction acts only on the vertical): compared with OPA results, EKE( $z$ ) profiles and eddy-driven features like the Zapiola anticyclone were much more realistic in SPEM simulations, even at  $1/3^\circ$  (de Miranda et al., 1999; Penduff et al., 2001).

The indexes defined in the present study were evaluated from a single simulation to show their physical relevance, highlight different aspects of this CLIPPER simulation, and suggest directions

for model development. Such metrics would be useful for model intercomparison as well. Depth-dependant validations of model outputs against CM datasets through similar metrics and through sensitivity studies would help better identify and correct the “baroclinic” discrepancies of different types of ocean models, further investigate the sensitivity of model solutions to numerics, assess the skill of other prognostic models and of operational models (such as those proposed in the MERSEA program). One may anticipate that future statistical studies, based on longer integrations of carefully-validated high-resolution models, may also help evaluate the “climatic” relevance of velocity and EKE statistics derived from intermittent current meter records (such as WOCE).

## Acknowledgments

Christel le Delliou and Didier Sorgniard are gratefully thanked for processing the model and WOCE data, and preliminary investigations. The authors would particularly like to thank Greg Holloway for interesting discussions and suggestions during his visit in LEGI. The CLIPPER project was supported by the Institut National des Sciences de l’Univers (INSU), the Institut Français de Recherche pour l’Exploitation de la Mer (IFREMER), the Service Hydrographique et Océanographique de la Marine (SHOM), and the Centre National d’Etudes Spatiales (CNES). Support for computations was provided by the Institut du Développement et des Ressources en Informatique Scientifique (IDRIS). Finally, the comments of the two anonymous reviewers are greatly appreciated.

## References

- Adcroft, A.J., Hill, C.N., Marshall, J., 1997. Representation of topography by shaved cells in a height coordinate ocean model. *Mon. Weather Rev.* 125, 2293–2315.
- Arhan, M., Colin de Verdière, A., Mercier, H., 1989. Direct observations of the mean circulation at 48°N in the Atlantic Ocean. *J. Phys. Oceanogr.* 19, 161–181.
- Barnier, B., 1998. Forcing the ocean. In: Chassignet, E.P., Verron, J. (Eds.), *Modeling and Parameterization*. Kluwer Academic Publishers, The Netherlands, pp. 45–80.
- Barnier, B., Reynaud, T., Beckmann, A., Böning, C., Molines, J.-M., Barnard, S., Jia, Y., 2001. On the seasonal variability and eddies in the North Brazil Current: insights from model intercomparison experiments. *Progr. Oceanogr.* 48, 195–230.
- Beismann, J.-O., Barnier, B., 2004. Variability of the meridional overturning circulation of the North Atlantic: sensitivity to the overflow of dense water masses. *Ocean Dyn.* 54 (1), 92–106.
- Blanke, B., Delecluse, P., 1993. Variability of the tropical ocean simulated by a general circulation model with two different mixed-layer physics. *J. Phys. Oceanogr.* 23, 1363–1388.
- Candela, J.S., Tanahara, S., Crepon, M., Barnier, B., 2003. The Yucatan channel flow: observations vs CLIPPER ATL6 and MERCATOR PAM05 models. *J. Geophys. Res.* 108 (C12), 3385.
- Chassignet, E.P., Arango, H., Dietrich, D., Ezer, T., Ghil, M., Haidvogel, D.B., Ma, C.-C., Mehra, A., Paiva, A.M., Sirkes, Z., 2000. DAMEE-NAB: The base experiments. *Dyn. Atmos. Oceans* 32, 155–183.
- Colin de Verdière, A., Mercier, H., Arhan, M., 1989. Mesoscale transition from the western to the eastern Atlantic along 48°N. *J. Phys. Oceanogr.* 19, 1149–1170.
- Crosnier, L., Barnier, B., Tréguier, A.-M., 2001. Aliasing of inertial oscillations in the 1/6° Atlantic Circulation Clipper model: impact on the mean meridional heat transport. *Ocean Modell.* 3, 21–32.

- de Miranda, A.P., Barnier, B., Dewar, W.K., 1999. On the dynamics of the Zapiola anticyclone. *J. Geophys. Res.* 104, 21,137–21,149.
- Fisher, N.I., Lee, A.J., 1983. A correlation coefficient for circular data. *Biometrika* 79, 159–166.
- Gille, S.T., Smith, S.G.L., 2000. Velocity probability density functions from altimetry. *J. Phys. Oceanogr.* 30, 125–136.
- Hall, N.M.J., Barnier, B., Penduff, T., Molines, J.M., 2004. Interannual variation of Gulf Stream heat transport in a high-resolution model forced by reanalysis data. *Clim. Dyn.* 23 (3–4), 341–351.
- Holloway, G., Sou, T., 1996. Measuring skill of a topographic stress parameterization in a large-scale ocean model. *J. Phys. Oceanogr.* 26, 1088–1092.
- Madec, G., Delecluse, P., Imbard, M., Lévy, C., 1998. OPA 8.1 Ocean General Circulation Model reference manual. Note du Pôle de Modélisation, Institut Pierre-Simon Laplace, No. 11, 91 pp.
- Maltrud, M.E., McClean, J.L., 2005. An eddy resolving global  $1/10^\circ$  ocean simulation. *Ocean Modell.* 8 (1–2), 31–54.
- Merryfield, W.J., Holloway, G., 1999. Eddy fluxes and topography in stratified quasi-geostrophic models. *J. Fluid Mech.* 380, 59–80.
- McClean, J.L., Semtner, A.J., Zlotnicki, V., 1997. Comparisons of mesoscale variability in the Semtner-Chervin  $1/4$  model, the Los Alamos Parallel Ocean Program  $1/6$  model, and TOPEX/POSEIDON data. *J. Geophys. Res.* 102, 25,203–25,226.
- McClean, J.L., Poulain, P.-M., Pelton, J.W., Maltrud, M.E., 2002. Eulerian and Lagrangian Statistics from surface drifters and a high-resolution POP simulation in the North Atlantic. *J. Phys. Oceanogr.* 32, 2472–2491.
- Pacanowski, R.C., Gnanadesikan, A., 1998. Transient response in a Z-level ocean model that resolves topography with partial cells. *Mon. Weather Rev.* 126 (12), 3248–3270.
- Penduff, T., Barnier, B., Béranger, K., Verron, J., 2001. Comparison of near-surface mean and eddy flows from two numerical models of the South Atlantic Ocean. *J. Geophys. Res.* 106, 16,857–16,867.
- Penduff, T., Barnier, B., Kerbiriou, M.-A., Verron, J., 2002. How topographic smoothing contributes to differences between the eddy flows simulated by sigma- and geopotential-coordinate models. *J. Phys. Oceanogr.* 32 (1), 122–137.
- Penduff, T., Barnier, B., Dewar, W.K., O'Brien, J.J., 2004. Dynamical response of the oceanic eddy field to the North Atlantic Oscillation: a model-data comparison. *J. Phys. Oceanogr.* 34 (12), 2615–2629.
- Reynaud, T., Legrand, P., Mercier, H., Barnier, B., 1998. A new analysis of hydrographic data in the Atlantic and its application to an inverse modelling study. *Int. WOCE Newsl.* 32, 29–31.
- Smith, W.H.F., Sandwell, D.T., 1997. Global seafloor topography from satellite altimetry and ship depth sounding. *Science* 277, 1957–1962.
- Stammer, D., Tokmakian, R., Semtner, A., Wunsch, C., 1996. How well does a  $1/4^\circ$  global circulation model simulate large-scale oceanic observations? *J. Geophys. Res.* 101, 25,779–25,811.
- Tokmakian, R., McClean, J., 2003. How realistic is the high frequency signal of a  $0.1$  degree resolution ocean model? *J. Geophys. Res.* 108, 3115.
- Tréguier, A.M., Reynaud, T., Pichevin, T., Barnier, B., Molines, J., de Miranda, A., Messenger, C., Beismann, O., Madec, G., Grima, N., Imbard, M., Le Provost, C., 1999. The CLIPPER project: high resolution modeling of the Atlantic. *Int. WOCE Newsl.* 36, 3–5.
- Tréguier, A.M., Barnier, B., de Miranda, A., Grima, N., Imbard, M., Le Provost, C., Madec, G., Messenger, C., Molines, J.M., Michel, S., Reynaud, T., 2001. An eddy permitting model of the Atlantic circulation: evaluating open boundary conditions. *J. Geophys. Res.* 106, 22,115–22,129.
- Tréguier, A.M., Boebel, O., Barnier, B., Madec, G., 2002. Agulhas eddy fluxes in a  $1/6^\circ$  Atlantic model. *Deep-Sea Res.*, II 50, 251–280.
- Tréguier, A.-M., Theetten, S., Chassignet, E., Penduff, T., Smith, R., Talley, L., Böning, C., Beismann, J.-O., in press. The North Atlantic subpolar gyre in four high resolution models. *J. Phys. Oceanogr.*
- Willebrand, J., Barnier, B., Böning, C., Dieterich, C., Killworth, P.D., Le Provost, C., Jia, Y., Molines, J.-M., New, A.L., 2001. Circulation characteristics in three eddy-permitting models of the North Atlantic. *Progr. Oceanogr.* 48 (2–3), 123–161.
- Woodgate, R.A., Fahrbach, E., Rohardt, G., 1999. Structure and transports of the East Greenland current at  $75^\circ\text{N}$  from moored current meters. *J. Geophys. Res.* 104, 18,059–18,072.