8-2016

Revisiting the Problem of the Gulf Stream Separation: On the Representation of Topography in Ocean Models with Different Types of Vertical Grids

Tal Ezer
Old Dominion University, tezer@odu.edu

Follow this and additional works at: https://digitalcommons.odu.edu/ccpo_pubs

Part of the Oceanography and Atmospheric Sciences and Meteorology Commons

Repository Citation
https://digitalcommons.odu.edu/ccpo_pubs/167

Original Publication Citation

This Article is brought to you for free and open access by the Center for Coastal Physical Oceanography at ODU Digital Commons. It has been accepted for inclusion in CCPO Publications by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
Revisiting the problem of the Gulf Stream separation: on the representation of topography in ocean models with different types of vertical grids

Tal Ezer

Center for Coastal Physical Oceanography, Old Dominion University, 4111 Monarch Way, Norfolk, VA 23508, USA

A B S T R A C T

The difficulty of simulating a realistic Gulf Stream (GS) that separates from the coast at Cape Hatteras has troubled numerical ocean modelers for a long time, and the problem is evident in different models, from the early models of the 1980s to the modern models of today. The source of the problem is not completely understood yet, since GS simulations are sensitive to many different factors, such as numerical parameterization, model grid, treatment of topography and forcing fields. A curious result of early models is that models with terrain-following vertical grids (e.g., “sigma” or “s” coordinates) seem to achieve a better GS separation than z-level models of similar resolution, so the impact of the vertical grid type on GS simulations is revisited here. An idealized generalized coordinate numerical model is used to compare between a sigma-coordinate grid and a z-level grid while maintaining the same numerical code and model parameters. Short-term diagnostic–prognostic calculations focus on the initial dynamic adjustment of the GS from a given initial condition and imposed boundary conditions. In diagnostic calculations, wherein the three-dimensional flow field is adjusted to time-invariant temperature and salinity data, the GS is quite realistic independent of the grid type. However, when switching to prognostic calculations, the GS in the z-level model tends to immediately develop an unrealistic GS branch that continues along the continental slope instead of separating from the coast at Cape Hatteras. The GS is more realistic in either a sigma-coordinate model or in a z-level model with a vertical wall replacing the continental slope. Increasing the vertical resolution in the z-level model reduces numerical noise, but it does not solve the GS separation problem. Vorticity balance analysis shows that the Joint Effect of Baroclinicity and bottom Relief (JEBAR) and its associated bottom pressure torque are very sensitive to the choice of vertical grid. A stepped topography grid may disrupt the local vorticity balance near steep slopes; this vorticity balance may be important to develop a counterclockwise circulation north of the GS that pushes the GS offshore. Therefore, the study suggests that a smooth representation of bottom topography in ocean models by using either a terrain-following coordinates or a z-level grid with partial cells may allow a more realistic treatment of flow–topography interactions and potentially a better simulation of the GS.

1. Introduction

The Gulf Stream (GS) is a western boundary current with a complex three-dimensional structure that is difficult to directly measure (e.g., Fuglister, 1963; Richardson and Knauss, 1971; Johns et al., 1995) and as difficult to realistically simulate with numerical models. One interesting aspect of the GS dynamics is that from the Florida Straits until Cape Hatteras it flows along the coast, but then it separates from the coast and turns farther eastward into the deep North Atlantic Ocean, rather than continue along the coast. Unfortunately, in many numerical models the simulated GS often tends to unrealistically loop toward the coast north of Cape Hatteras, and separates from the coast farther north than observed (Bryan and Holland, 1989; Semtner and Chervin, 1988; Thompson and Schmitz, 1989; Chassignet et al., 2003; Schoonover et al., 2016). Attempts to study the “Gulf Stream separation” issue started early on with simple idealized models that show, for example, the important role of wind and stratification on model results (Parsons, 1969; Nurser and Williams, 1990). Other early models with an idealized topography and a simplified vertical structure such as barotropic models (e.g., Deng, 1993) or quasigeostrophic models (e.g., Özgökmen et al., 1997) evaluated the role of wind, eddies, the shape of the coastline, the GS inertia and the slip/no-slip model boundary conditions. Primitive equations models with an idealized topography were also used to demonstrate the impact of the Deep

E-mail address: tezer@odu.edu

http://dx.doi.org/10.1016/j.ocemod.2016.05.008
Western Boundary Current (DWBC) and recirculation on GS separation (e.g., Spall, 1996). How much can be learned from these idealized models about the real GS is questionable, given, for example, the fact that some early models represented the continental slope by a vertical wall and neglected the coastal ocean. Therefore, studies of the GS separation were extended to coastal ocean models; these models show for example that to achieve a realistic GS separation, models may need to include local surface heat flux over the shelf and need to resolve the recirculation gyre between the GS and the coast (Ezer and Mellor, 1992). In the late 1980s models’ resolution became fine enough to resolve the GS front and mesoscale eddies, at least to some degree and to include more realistic topography and coastline. Nevertheless, unrealistic GS separation has been a lingering problem in many models even today, though simulations do improve when very high horizontal resolution is used (Smith et al., 2000; Bryan et al., 2007; Chassignet et al., 2008; Hurlburt and Hogan, 2000, 2008; Hurlburt et al., 2011; Schoonover et al., 2016).

Two secondary problems of unrealistic GS path in ocean models include: (1) simulated temperatures in the Mid-Atlantic Bight may be warmer by several degrees than observed, causing problems in coupled ocean–atmosphere models, and (2) the southward flowing cold Slope Current (Rossby et al., 2010) may be missing or be too weak, and the northern recirculation gyre north of the GS (Mellor et al., 1982; Hogg, 1992) is thus not well simulated. The two aspects above are especially important for climate modeling. For example, a recent study (Saba et al., 2016) demonstrates how a misplaced GS in coarse resolution climate models affect climate simulations, so that a higher resolution ocean and atmospheric models with more realistic GS representation results in enhanced warming in the northwest Atlantic Ocean in future climate change simulations. Moreover, recent studies connect climate-related variations in the GS to coastal sea level rise and increased flooding along the U.S. East Coast (Ezer et al., 2013; Yin and Goddard, 2013; Sweet and Park, 2014; Ezer, 2015), thus reemphasizing the need of climate models to more accurately represent the GS, if coastal sea level rise is to be accurately predicted.

The source of the GS separation problem in ocean models is still not completely understood since a model’s GS depends on so many different factors such as surface forcing (Ezer and Mellor, 1994), model coastline (Deng, 1993), wind and eddies (Özgökmen et al., 1997), grid resolution (Hurlburt and Hogan, 2000), boundary conditions (Thompson and Schmitz, 1989; Ezer and Mellor, 1994, 2000 ), eddy-driven abyssal circulation and DWBC (Hurlburt and Hogan, 2008) and various numerical aspects such as subgrid-scale parameterizations (Chassignet and Garraffo, 2001; Chassignet et al., 2003; Chassignet and Marshall, 2008; Schoonover et al., 2016). It is thus likely that the GS separation in each model is the result of not one factor, but a combination of several factors mentioned above. One of the factors that could significantly affect the GS separation in ocean models is the way bottom topography is represented by the model grid – this can influence the flow–topography interaction. For example, Myers et al. (1996) found that in ocean models the bottom pressure torque component of the Joint Effect of Baroclinicity and bottom Relief (JEBAR) was significantly different than that obtained directly by diagnostic calculations, and that the JEBAR term is crucial for the GS separation. The JEBAR may influence the flow in regions where vertical stratification and bottom slopes interact (for detailed discussions of the role of JEBAR in ocean models see Sarkisyan and Ivanov, 1971; Mellor et al., 1982; Greatbatch et al., 1991; Cane et al., 1998; Sarkisyan, 2006; Xu and Oey, 2011, and many others). The role of the bottom pressure torque in GS dynamics was also addressed in a recent study (Schoonover et al., 2016), suggesting that the GS separation is related to local dynamics rather than to the wind-driven basin-scale dynamics. The implication is that local flow–topography interac-

tions may be important, but they may not be accurately simulated in some models. A curious related result in early simulations is that given the same moderate horizontal grid resolution (~20 km), GS separation is more realistic in models with smooth representation of topography, such as in models with terrain-following (e.g., sigma or s coordinates) vertical grids (Ezer and Mellor, 1992, 1994, 1997, 2000; Ezer, 1999; Haidvogel et al., 2000) than in models of similar resolution that use step-like z-level vertical grids (Bryan and Holland, 1989; Semtner and Chervin, 1988). Early models of the Atlantic Ocean using the Hybrid Coordinate Ocean Model (HYCOM) also show some deficiencies in GS simulations (Chassignet et al., 2003). The recent model intercomparison study of Schoonover et al. (2016) confirms the early results, by showing that the GS separation is quite realistic in a terrain-following model (the Regional Ocean Modeling System, ROMS) and in a model with partial cell representation of bottom topography (the MIT general circulation model, MITgcm), compared with an unrealistic northern GS separation in a z-level model (the Parallel Ocean Program, POP). However, the above study could not attribute the differences in GS separation to model grid types, because the models in the study use different numerical schemes, different subgrid-scale parameterizations and different horizontal grid sizes (POP, 10 km and 100 km; MITgcm, 3 km and 10 km; and ROMS, 2.5 km and 6 km).

The advantage of smooth representation of topography in sigma models (or other terrain-following models) is contrasted with the potential disadvantage of sigma models with regard to numerical errors associated with the pressure gradient term over steep topography (Mellor et al., 1998; Ezer et al., 2002). For the fine grid resolution and smooth topography of the sigma coordinate model used here, the numerical errors associated with pressure gradient errors were found to be small (order of mm s\(^{-1}\)) compared with the mean flow and other errors. The hypothesis that the different representation of bottom topography in z-level and in sigma models impact the GS separation is difficult to test, because different models often use very different numerical schemes and mixing parameterizations, so model-to-model inter-comparison studies (Willems et al., 1994; Chassignet et al., 2000; Ezer et al., 2002; Schoonover et al., 2016) cannot isolate the influence of the choice of vertical coordinate from among the other differences between models. A solution is to use a generalized-coordinate ocean model in which one can apply exactly the same model setup and numerical schemes except the vertical grid. Such comparisons of z-level and sigma models indeed show large sensitivity to vertical grid type in simulations of wind-driven ocean circulation (Mellor et al., 2002), in simulations of deep water formation (Ezer and Mellor, 2004) and in simulations of dense overflows (Ezer, 2005, 2006). Therefore, the same generalized-coordinate model developed by Mellor et al. (2002) (which is based on the Princeton Ocean Model, POM) will be used here. The main goal of the study is to test the hypothesis that the representation of topography in ocean models can strongly affect the GS separation, and if true to find the mechanism involved. Benefits of such a study are twofold: first, to get a better understanding of numerical ocean models behaviors and the dependence of that behavior on the user’s choices of grids, and second, to get a better understanding of the processes that control the GS dynamics and its interaction with topography.

The paper is organized as follows. First, the numerical model setup and the different experiments are described in Section 2, then a comparison of the results of different simulations are described in Section 3, following by analysis of the dynamical balances in Section 4. Finally, a summary and conclusions are offered in Section 5.
2. Numerical model setup and experiments

The model is based on the generalized coordinate numerical ocean circulation model of Mellor et al. (2002), which includes a Mellor–Yamada turbulence scheme and Smagorinsky-type horizontal diffusion. With one switch, the same numerical scheme can be used as a sigma-coordinates model (e.g., Ezer and Mellor, 1992), a z-level model (e.g., Bryan and Holland, 1989), an s-coordinates model like ROMS (Haidvogel et al., 2000) or a combined sigma/z-level model (Ezer and Mellor, 2004). The model domain and configuration used here is the same as in Ezer (2016), Ezer (2016) and this study use an idealized model topography, which is a useful way to investigate flow–topography interaction processes, as previously done with the same numerical model (Ezer and Mellor, 2004; Ezer, 2005, 2006). The model topography is a smoothed version of the real coastline and bathymetry (Fig. 1), neglecting rivers, estuaries, barrier islands, seamounts, etc. This topography is interpolated and discretized for each model grid. The minimum depth is set to 10 m and the maximum depth is set to 3000 m, focusing on the interaction of the GS with the continental shelf and slope; this allows a longer barotropic time step (for the vertically integrated equations) than a deeper domain will require. Barotropic and baroclinic time-steps of 17 s and 8 min, respectively, were used for the sigma model, but a smaller baroclinic time-step of 3 min was needed for the z-level model to suppress noise on the shallow shelf region (which is not well resolved with a z-level grid). The model is driven at the surface by a constant monthly mean wind (May 2012; see Ezer, 2016), but for the short-term simulations done here the wind has negligible impact, as shown later by the dynamic balance analysis. Surface heat and freshwater fluxes are set to zero. Though heat flux may play a role in getting a realistic GS separation in long-term simulations, as shown by Ezer and Mellor (1992), heat flux can be neglected for the very short-term idealized simulations conducted here. Inflow/outflow transports are imposed on the eastern and southern open boundaries (Fig. 1b shows those transports). The horizontal grid is a Cartesian grid with 1/12 resolution (~6 to 8 km grid size) and the vertical sigma grid has 21 layers with a higher resolution near the surface (e.g., the thickness of each layer varies from ~1/1000th to 1/15th of the water depth between the surface and bottom layers). The basic z-level grid has exactly the same vertical resolution as the sigma grid in the deepest part of the domain, but a more coarse vertical resolution in shallow regions (in z-level coordinates the top layers above the bottom are active while deeper layers represent land). Note that this vertical resolution is somewhat coarse compared with modern z-level models, so another experiment with 61 layers is also conducted; this higher vertical resolution resemble the vertical grid in the POP model used by Schoonover et al. (2016). The model domain and its boundary inflow/outflow conditions are very similar to the early regional GS models of Mellor and Ezer (1991) and Ezer and Mellor (1992); this model differs from the previous model by having higher resolution, smoother coastline, and the focus on the short-term dynamic adjustment process. Using an idealized topography helps to isolate the impact of the basic topographic features of the region on the GS. The recirculation gyres north and south of the GS are important parts of the GS dynamics, as seen in diagnostic calculations of the Atlantic Ocean circulation (Mellor et al., 1982; Ezer and Mellor, 1994), thus regional models must include these gyres in their boundary conditions to obtain a realistic GS, as demonstrated by Ezer and Mellor (1992). Here, three inflow transports are imposed: the Florida Current (FC), the Slope Current (SC) and the Sargasso Sea (SS) and their total transport is equal to the outflow of the Gulf Stream (GS), as seen in Fig. 2b. Only the total transport (vertically integrated velocity) is specified on the boundary together with standard barotropic radiation boundary conditions to minimize artificial reflection of waves from the boundary. The vertical distribution of the velocity near the open boundaries is calculated by the model from the density field in a 1° buffer zone near the southern and eastern open boundaries. The transport and location of the barotropic inflow/outflow on the boundaries are fixed for the idealized simulations performed here. The impact of time-dependent transports in the same model was the topic of the study of Ezer (2016). Note that using a barotropic inflow/outflow conditions on the eastern boundary means that the DWBC is not specifically imposed, so its role in the GS separation (Spall, 1996; Thompson and Schmitz, 1989; Hurlburt and Hogan, 2008) has not been assessed here.

Initial condition is the monthly mean temperature and salinity field obtained from reanalysis data (Ferry et al. 2012) for May 2012. The data are interpolated from 1/4° grid and 33 vertical layers into the model grid. Simulations (not shown) indicate that the

![Fig. 1. (a) Bottom topography (color, in m) of the region and schematics of the main currents. (b) The model domain, its simplified topography and the location of inflow/outflow boundary conditions. This is the topography used in experiments SIG and 21V (see text for details on experiment ZNS). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image-url)
basic results are not much affected by the choice of the month since they are focused on an idealized short-term adjustment process. Note however, that for long-term simulations, one may need to add time-varying winds, freshwater and surface heat fluxes, etc., as has been done in more realistic models of the region that are used for process studies (Xu and Oey, 2011) or operational forecast systems (Aikman et al. 1996); conducting realistic simulations is beyond the scope of this study.

The main focus of the study is on the diagnostic–prognostic adjustment problem, following previous studies of that nature (Ezer and Mellor, 1994). In the first stage, starting from a zero velocity field, a 10-day diagnostic calculation is conducted, where the temperature and salinity fields are held unchanged and equal to the initial condition (i.e., neglecting advection and diffusion terms in the heat and salt equations); this calculation develops the three-dimensional flow field associated with the initial density. Then, in the prognostic stage, tracer properties are allowed to change and a new semi steady state is reached. The prognostic adjustment calculation is relatively short (50 days in this case), and its main purpose is to allow the model to adjust the density field in order to obtain a more dynamically balanced state. Much can be learned about models and processes from the way the fields are dynamically adjusted. Ezer and Mellor (1994) demonstrated that the results of the diagnostic–prognostic simulations produce a very similar circulation pattern as that obtained by pure diagnostic models using the same density and wind data (Mellor et al., 1982; Greatbatch et al., 1991). These types of simulations are useful for process studies, not for representing long-term variability which requires time-dependent forcing over many years. Because of the small domain, the constant forcing and the strong influence of the imposed boundary conditions, the three-dimensional velocity field and the surface elevation in the model are dynamically adjusted to the density field very quickly, as will be shown by the average kinetic energy.

In the general coordinate model (Mellor et al., 2002) the vertical grid transformation is \( z = \eta(x, y, t) + s(x, y, k, t) \), where \( \eta \) is the surface elevation, \( s \) is the vertical grid distribution, and \( k \) is the vertical layer number \( (k=1,2,...,M, \text{for M layers}) \); for different grids users can specify different functions for \( s \) (e.g., the \( s \)-coordinates in ROMS is a specific case). Four different simulations are conducted with the same forcing and initial conditions; they differ only by the model vertical grid and the bottom topography.

1. Experiment “SIG” is a simulation with a sigma coordinates model, where the vertical grid has \( M = 21 \) layers and the distribution is \( s = \sigma(k)[H(x,y) + \eta(x,y,t)]; \ -1 < \sigma < 0; \ -H < z < \eta. \)
2. Experiment “ZLV” is a simulation with a z-level model, where the vertical grid has \( M = 21 \) layers and the distribution is \( s = \sigma(k)[H_{max} + \eta(x,y,t)]. \ H_{max} = 3000 \text{ m} \) (same \( \sigma \) as in SIG).
3. Experiment “Z60” is a z-level simulation as ZLV, but with a higher vertical grid, \( M = 61 \). In this case the vertical grid size varies between \( -1 \text{ m} \) near the surface to \( -50 \text{ m} \) in the deepest regions.
4. Experiment “ZNS” (z-level no slope) is a simulation with the same grid as in ZLV, but with no continental slope north of Cape Hatteras. In this case, for latitudes \( > 35^\circ \text{N} \) and \( H > 100 \text{ m} \) the continental slope is replaced by an almost a vertical wall (i.e., depth drops immediately from 100 m to 3000 m).

Case 1 represents a terrain-following grid, while cases 2–4 represent z-level grids with almost horizontal fixed layers. Note that with free surface, the vertical grid is actually spatial- and time-dependent (sometimes called a \( z^* \) grid; Adcroft and Campin, 2004), but the deviation from horizontal layers is assumed to be very small \( (\eta(x,y,t)<H_{max}) \), so no correction to pressure gradient is applied here with regard to the ZLV calculations. In the ZLV, Z60 and ZNS experiments the topography is thus represented by steps (ZLV with large steps, Z60 with smaller steps and ZNS with mostly one giant step in the northern part of the domain). It should be acknowledged that some z-level models try to improve the representation of topography by using shaved or partial cells.
(Adcroft et al., 1997; Pacanowski and Gnanadesikan, 1998). However, the experiments here are based on the basic step-like z-level grid that has been used in the early models that had troubles simulating realistic GS separation, as discussed before. It is noted that the GS separation in a z-level model with partial cells may look more like the results of a terrain-following model (s- or sigma-coordinates) than that of a stepped-topography z-level model, according to the experiments presented by Schoonover et al. (2016).

3. Comparisons between the sigma-coordinates and the z-level models

The dynamic adjustment in the basic experiments during the diagnostic–prognostic calculations can be seen in the mean kinetic energy (MKE; Fig. 2), calculated from the vertically averaged velocity over two sub-regions, the latitudes of the Mid-Atlantic Bight (MAB, north of 36°N) and the latitudes of the South Atlantic Bight (SAB, south of 36°N). The vertically averaged velocity is used here to be consistent with the vertically averaged dynamics analyzed later (evolution of the MKE calculated from the three-dimensional velocity field would yield a similar pattern). Starting from an initial condition of a motionless ocean, it takes about 2–3 days for velocities to develop during the diagnostic run; in this stage the flow field is driven mainly by density gradients associated with the initial temperature and salinity field (and to lesser degree by wind). When switching to a prognostic run after 10 days, the MKE is initially reduced within another 2–3 days, as diffusion and advection smooth out noisy data (Ezer and Mellor, 1994). During the following days the flow field is adjusted to the bottom topography, until a new semi-steady state is reached (or not). The results show large differences between the SAB and the NAB regions. In the south, when the GS is still flowing close to the coast, the ZLV and SIG-experiments are quite similar (with somewhat larger MKE for the ZLV case). However, in the north, after the GS passed Cape Hatteras, where it does in reality separate from the coast, there are significant differences between the experiments, with a much higher MKE in the ZLV case, compared with the SIG results (the reason for this difference will be discussed in details later).

The adjustment process in the two basic experiments is demonstrated in an east–west velocity cross-section at 35°N (near Cape Hatteras) for the ZLV and SIG experiments (Fig. 3a–d). The pure diagnostic calculations (Fig. 3a and c) show narrow and deep northward flowing GS (down to over 2000 m) with strong currents all the way to the bottom itself; this flow pattern is not so realistic when compared with observations at that region (Fig. 3e; after Richardson and Knauss, 1971). On the other hand, after the prognostic calculations (Fig. 3b and d) the flow pattern seems much more realistic, resembling the observed GS (Fig. 3e) in its width (~100 km), depth (0.2 m s⁻¹ contour reached ~1000 m) and the eastward tilt with depth. The observed southward flowing return current both east and under the GS (part of the DWBC?) is obtained in the two models, but the deep flow is stronger in the ZLV case than the SIG case. This result demonstrates the usefulness of the diagnostic–prognostic approach in reconstructing three-dimensional flow field from hydrographic observations. The original data itself most likely had errors and thus required adjustment to produce a dynamically balanced flow field. A notable difference between the SIG and ZLV runs is that bottom boundary layers are better resolved with the sigma coordinates, resulting in a smoother flow that decays toward the bottom (Fig. 3b) compared with a more noisy near bottom flow field in the z-level model (Fig. 3d).

The sea surface height (SSH) fields at the end of the diagnostic runs (Fig. 4) indicate very similar results for all four experiments. An artificially high sea level on the shelf at ~39°N is seen only in the ZNS case (Fig. 4c), which is likely due to the sudden drop in topography in this simulation (white lines represent topography contours). The main path of the GS is consistent with observations and simulated quite well by all models. Even the signature of two cold-core eddies (around 30°N and 35°N) were similarly simulated by all runs. In contrast with the consistent results of the four diagnostic calculations (Fig. 4), during the prognostic runs the GS in the MAB evolved very differently in the four cases (Figs. 5 and 6). After 30 days, the GS separation point in SIG changed very little (Fig. 5a; slightly moved north), it moved south in ZNS (Fig. 5c), but moved significantly north (to ~39°N) in ZLV and Z60 (Fig. 5b and d). The similarity between ZLV and Z60 is the first indication that the resolution of the vertical grid alone does not seem to affect the GS separation. After 60 days (Fig. 6), the results of the four cases further separated from each other. In all cases the GS tend to overshoot Cape Hatteras to some degree, but the largest departure from the observed GS was in ZLV and Z60, which are the z-level cases with realistic slopes (Fig. 3d); in these cases the GS and eddies tend to continue moving northeastward along the shelf break. Somewhat surprising is the result obtained by the z-level model with unrealistic continental slope (ZNS; Fig. 6c). In both, SIG and ZNS cases a recirculation gyre was developed north of the GS, which helped to move the GS to a more realistic path away from the coast. As mentioned before, several past studies indicate the importance of this gyre for GS separation (e.g., Ezer and Mellor, 1994). The results should not be interpreted as implying that all z-level models will produce an unrealistic GS like that in Fig. 6b, in fact, long-term simulations with a different treatment of topography (say partial cells), time-dependent wind and heat flux forcing and a larger domain may eventually restore the GS into a more realistic path. Though most modeling studies only show the results after a lengthy spin-up (several weeks or months for small regional models and up to many years in large-scale models), it is constructive to understand why the initial tendency of the GS to separate or not is so sensitive to the choice of vertical grid.

An example of a north–south velocity cross-section after 60 days at 70°W demonstrates the large differences between the models (Fig. 7). The GS in the SIG case (Fig. 7a) is quite realistic, though the flow is a little more diffused; the GS also turns southeast at that longitude so the shown eastward u-velocity component is smaller than the along-stream velocity. Because of the idealized topography and imposed barotropic inflow in the northeast, deep currents are weaker than in realistic sigma-coordinate basin scale models that show very intense DWBC (Ezer and Mellor, 1997). In the two z-level cases with realistic continental slope the GS seems to split into an offshore branch and another unrealistic branch of eastward flowing slope current (around 39°–40°N) – this is the current that flows along the coast all the way from Cape Hatteras to the eastern boundary of the model. This unrealistic eastward flowing (red) warm current pushes the westward flowing (blue) cold current away from the continental slope. The removal of the continental slope in the ZNS experiment seems to eliminate this unrealistic current and instead a more robust westward flowing current is developed between the offshore GS and the coast, keeping the GS away from the coast (Fig. 7c). The westward flowing (blue) return flow is not as strong in SIG as it is in ZNS (it is probably also less defined than observed). Since a barotropic inflow boundary condition is imposed on the eastern boundary at 65°W, a realistic SC and DWBC are not expected in these idealized simulations. These experiments suggest that the representation of the continental slope in ocean models may strongly affect the GS separation, so further analysis of the dynamics involved is conducted next.

4. Dynamic balance analysis

To further investigate how the vertical grid may affect the dynamics in those experiments, a dynamic balance analysis is
conducted using the barotropic vorticity equation. The formulation follows a similar analysis that was conducted in previous studies of bottom boundary layers (Ezer and Mellor, 1994; Ezer, 2005) and in studies of the Atlantic Ocean circulation (Ezer and Mellor, 2000). Several other studies used this approach to study the dynamics in ocean models (e.g., Schoonover et al., 2016), though they may use a slightly different formulation for the same equation. The vertically integrated vorticity balance equation can be written as,

\[
\frac{\partial}{\partial t} \left( \frac{\partial V D}{\partial x} - \frac{\partial U D}{\partial y} \right) + \left( \frac{\partial A y}{\partial x} - \frac{\partial A x}{\partial y} \right) + \frac{\partial (f U D)}{\partial x} + \frac{\partial (f V D)}{\partial y} = \left( \frac{\partial P b}{\partial x} \frac{\partial D}{\partial y} - \frac{\partial P b}{\partial y} \frac{\partial D}{\partial x} \right) + \left( \frac{\partial r y}{\partial x} \frac{\partial r x}{\partial y} \right) - \left( \frac{\partial r y b}{\partial x} \frac{\partial r x b}{\partial y} \right)
\]  

\(1\)
Fig. 4. Sea surface height anomaly fields (SSH; colorbar in m) at the end of the diagnostic calculations (10 days). Results are shown for experiments: (a) SIG, (b) ZLV, (c) ZNS and (d) Z60. Bathymetry contours for 100, 1000 and 2000 m are shown in white lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 5. Same as Fig. 4, but for the prognostic runs at day 30. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)
Fig. 6. Same as Fig. 4, but for the prognostic runs at day 60. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Fig. 7. East–west velocity (u-component in m s⁻¹) across 70°W after 60 days for experiments: (a) SIG, (b) ZLV, (c) ZNS and (d) Z60. Contour intervals are as in Fig. 3. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)
where \((U, V)\) are the vertically averaged velocity components, \(D = H + \eta\) is the water depth plus surface elevation, \((\varepsilon_U, \varepsilon_V)\) are the advection and diffusion terms (here advection>>diffusion), \(P_b\) is the bottom pressure, \((\tau_{\varepsilon_U}, \tau_{\varepsilon_V})\) are the surface friction stress components and \((\tau_{\varepsilon_b}, \tau_{\varepsilon_D})\) are the bottom friction stress components. The six terms in (1) represent the barotropic vorticity tendency (time dependent) term, the non-linear advective and diffusive torques, the Coriolis term, the bottom pressure torque, the curl of the surface stress and the curl of the bottom stress. Note that the bottom pressure torque includes two contributions to pressure (Eq. (2)) – the first one is the surface elevation term that involves surface elevation gradients and drive barotropic flows, and the second one involves baroclinic flows and its contribution to the bottom pressure torque is the so-called JEBAR (Mellor et al., 1982; Greatbatch et al., 1991; Sarkisyan, 2006).

Fig. 8 shows an example of the leading terms in (1) in a north–south section at 70°W after 30 days (same section as in Fig. 7) for the four experiments (note the different y-scale for each experiment). The two terms of the bottom pressure torque are shown separately; they usually have opposite signs, but they do not completely cancel each other, as discussed later. In the experiments conducted here, the surface and bottom stress terms are generally of an order of magnitude smaller than the other terms so they are not shown. The advective–diffusive term is dominated by the non-linear advection so for convenience this term is marked here as advection. It is immediately clear from Fig. 8 that the choice of model grid has a very significant impact on the barotropic vorticity in the model. In the sigma model (Fig. 8a), the largest terms are near the GS region (around 38°N), where the advection and tendency are the dominant terms, representing the meandering GS and eddies with their large velocity gradients. On the continental slope (39–40°N) the surface elevation (blue) and the JEBAR (green) terms seem to balance each other (so the net bottom pressure torque is small). However, near the shelfbreak, in addition to the bottom pressure torque, the Coriolis and the advection are also needed for obtaining a balanced equation. The results are quite consistent with an area averaged integrated vorticity analysis near the GS that show the dominant terms in different models to be the bottom pressure torque and the Coriolis term (Schoonover et al., 2016). The important role of JEBAR in the dynamics of flows over the continental slope and shelfbreak in this region was also the topic of detailed analysis by Xu and Oey (2011). The results from the coarse vertical resolution z-level model (Fig. 8b) are very different than the sigma-model results, especially along the slope and shelf, where the terms are ~5 times larger in the ZLV experiment than those of the SIG experiment. Schoonover et al. (2016) also found larger vorticity terms in a coarse resolution z-level model (POP) than in a terrain-following model (ROMS) or in a higher resolution z-level model (their Fig. 5), but they did not specifically discussed the impact of the vertical grid type. Large “see-saw” like spatial oscillations between 39°N and 40°N shows the JEBAR and surface elevation gradients terms with opposite signs– these are the only terms that involve gradients of bottom depth, and the peaks are located where the z-level steps are found (Fig. 7d). Near the unrealistic coastal branch of the GS (~39.7°N) a maximum positive JEBAR peak is balanced by the negative advection and surface elevation gradient terms. When the vertical resolution is tripled (Z60 case) the “see-saw” noise over the slope disappears, but the main balance between JEBAR, surface elevation, advection and tendency remains (Fig. c), indicating a fundamental impact of z-level grid on vorticity. The simple replacement of the continental slope in the ZLV case with one vertical wall in the ZNS case, eliminates all the large peaks except at the shelfbreak (~40.2°N; Fig. 8d) where the topography suddenly changes in this case. Note that, in the open ocean away from the coast, all three experiments have a similar advection-tendency balance with comparable amplitudes (though owing to the different vertical scale, this may not be so clear in Fig. 8). It is clear from these experiments that the representation of the continental slope in the models and the resultant JEBAR dominate the differences between the three experiments.

The two bottom pressure torque (BPT) terms (the barotropic and baroclinic parts in 2) are the only terms that depend on the bottom slope – they seem to almost cancel each other (Fig. 8), but small imbalances between them may still exist (especially over the shelf break in the coarse vertical grid z-level; Fig. 8b). Therefore, the sum of the two terms (i.e., net bottom pressure torque) is averaged over the continental slope at (70°W, 39–41°N) and over the first 10 days of the prognostic runs. The results show striking differences in the mean net value of the BPT (in units of 10^{-10} m^2 s^{-2}) for the 4 experiments: SIG = +3.5, ZLV = −147, Z60 = −55, ZNS = −4.2. While the actual values in this example are not important, several conclusions can be drawn. First, a positive BPT in SIG can drive a positive vorticity tendency (a counter-clockwise circulation as in the northern recirculation gyre), while a negative BPT in all the three z-level cases can drive a clockwise circulation pattern (i.e., as seen when the GS continues unrealistically along the coast). Second, either increasing the vertical resolution (Z60) or removing the slope (ZNS) can significantly reduce the imbalance between the two BPT terms in z-level models. It is also noted that the mean non-linear advective–diffusive terms over the slope are larger in all the z-level cases than they are in SIG, indicating that the stepped topography is associated with noisier near-bottom velocities than models with smooth representation of the bottom; the latter is consistent with similar findings in early models (Mellor et al., 2002).

The initial adjustment process in terms of the absolute value of the barotropic vorticity is averaged along the same section as in Fig. 8 and shown as a function of time in Fig. 9. First, it is noted that the time-scale of the initial adjustment, both, in the diagnostic and in the prognostic runs are very short, order of only few days. The fact that the amplitudes of the tendency and advection terms are reduced during the prognostic run from the higher level of the diagnostic run and remain almost unchanged indicate that a reasonable dynamic adjustment is achieved. Keep in mind that during the diagnostic run, advection of tracers is zero (density is kept unchanged), but the non-linear velocity advection terms may play an important role in the dynamic adjustment. Second, a substantial similarity is found in the mean amplitude of the terms in the four experiments during the adjustment (Fig. 9), despite large differences in the local dynamics (Fig. 8). (Note that during the diagnostic run in the ZNS case the one huge drop in the continental slope dominates the JEBAR and surface elevation terms, but eventually these terms almost completely cancel each other during the prognostic run, as discussed before). These findings are consistent with the results of Schoonover et al. (2016) that show that the area-integrated vorticity budget is similar across different types of models and different resolutions (as in Fig. 9), while the GS separation may be more closely related to local small-scale dynamics that is associated with flow–topography interactions over continental slopes (as in Fig. 8).

5. Summary and conclusions

The problem of the Gulf Stream separation in ocean models has troubled ocean modelers for almost 3 decades (Semtner and Chervin, 1988; Bryan and Holland, 1989; Thompson and Schmitz, 1989; Ezer and Mellor, 1992; Deng, 1993; Myers et al., 1996; Özgökmen et al., 1997; Chassignet and Garraffo, 2001; Chassignet et al., 2003; Bryan et al., 2007; Chassignet and Marshall, 2008;
realistic GS separation, models need to resolve not only the GS itself (which require high enough horizontal resolution), but also the southward flowing slope and shelf currents that are the northern branches of the recirculation gyre north of the GS, as seen in models (Mellor et al., 1982; Greatbatch et al., 1991; Ezer and Mellor, 1994; Bryan et al., 2007) and observations (Hogg, 1992; Rossby et al., 2010). The region of interest has large variations in stratification and steep continental slopes, so the flow–topography interaction is often assessed through the JEBAR term (Mellor et al., 1982; Greatbatch et al., 1991; Myers et al., 1996; Sarkisyan, 2006).

The focus of this study is on the representation of topography in ocean models, and in particular, on the role of vertical grid type. Anecdotal evidence that model grid type affects GS
Fig. 9. The time evolution of the amplitude (absolute value) of the leading terms of the vorticity balance equation during the first 20 days of the adjustment, averaged along the same 70°W section as in Fig. 7 and for experiments: (a) SIG, (b) ZLV, (c) Z60 and (d) ZNS.

separation can be seen in past basin-scale and global models, whereas the GS seems more realistic in models that use terrain-following grids (e.g., Ezer and Mellor, 1997, 2000; Haidvogel et al., 2000) and less realistic in models that use z-level grids (Semtner and Chervin, 1988; Bryan and Holland, 1989), even though all these models had similar moderate resolutions (~20 to 50 km). This motivates this study to conduct a more systematic test that will address the impact of model grid type. The treatment of topography in z-level models can be improved with schemes such as partial or shaved cells (Adcroft et al., 1997; Pacanowski and Gnanadesikan, 1998), but such improvements are not addressed here. Also noted is the fact that when the horizontal and vertical resolution in a z-level model is significant increased, the solution of near-bottom flows is converged to the solution of a sigma-coordinate model with a coarser resolution (Ezer and Mellor, 2004). The present study compares a sigma-coordinate model with the simplest z-level model with stepped topography. The simplicity of the model configuration will emphasize the differences, since there is no attempt here to optimize the grids.

The experiments focus on the initial stages of the dynamic adjustment process using short-term diagnostic–prognostic calculations with a generalized-coordinates ocean model using an idealized smooth topography and GS that is driven by imposed boundary conditions. Under these control conditions it was easy to
detect the impact of the grid type and topography, when all other numerical elements are identical. In contrast, model intercomparison studies that show significant differences in GS separation between different models (e.g., Schoonover et al., 2016) could not isolate the impact of the choice of vertical grid, because several other factors in the models were different. The tendency of the GS in the z-level model to continue flowing along the coast of the MAB instead of separating at Cape Hatteras was immediately clear within a few days of switching from diagnostic to prognostic calculations. This unrealistic GS separation is consistent with some early z-level models (e.g., in Bryan and Holland, 1989, the GS separated from the coast at ~40°N to 42°N instead of ~35°N). The results here seem even somewhat worse than in those basin-scale models, since in long-term simulations with larger domains and realistic surface forcing, the GS may “pushed” to eventually separate from the coast (though often at northern latitude than observed). However, the comparison with the sigma-coordinate model clearly demonstrates that the vertical model grid type does have a crucial impact on the GS separation. The source of the problem seems to be the representation of the continental slope by the z-level stepped-topography grid, as demonstrated by a special z-level model experiment without continental slope (just a vertical wall). In this special case, the z-level model results resemble the GS obtained by the sigma-coordinate (and the observed GS) more than it does the standard z-level model. Increasing the vertical resolution in the z-level reduces noise over the slope, but did not result in a realistic GS separation, thus pointing to a fundamental issue with a stepped-topography representation in z-level models. Further analysis of the dynamic balance across the continental slope and the GS reveals that the source of the discrepancy between the sigma and the z-level models is in fact due to the representation of slopes in the model. If the vertical resolution in the z-level is insufficient, the stepped-topography creates spikes in the two terms that composed the bottom pressure torque. Near the shelf break there is a balance between the JEVAR term, the surface elevation gradient term and the advection term. However, the amplitude of these terms in the z-level model depends on the size of the topographic steps in the model rather than on the real dynamics. The role of flow–topography interaction as presented by the JEVAR plays a role in the process of GS separation, as previously suggested (Myers et al., 1996), so that a smoother representation of slopes as done in terrain-following models or in models with partial cells may be important to accurately represent this term in the model. The results are also consistent with the findings of Schoonover et al. (2016) that show that GS separation is affected more by local dynamics of the bottom pressure torque rather than by large-scale wind-driven balance. In the experiments conducted here, the local net bottom pressure torque over the continental slope indicated the generation of barotropic vorticity that was positive in the sigma-coordinate model and negative in the z-level models, the former will drive a more counterclockwise circulation pattern (as in the observed northern recirculation gyre) and the latter will drive a more clockwise circulation pattern (i.e., with an unrealistic northeastward flow along the slope).

In summary, it was demonstrated here that the choice of vertical grid type does have a significant impact on the dynamics of the GS. A better representation of topographic slopes in all types of ocean models is important not only for processes such as bottom boundary layers, overflow dynamics and topographic waves, but also for western boundary currents that affect large-scale ocean circulation.

Acknowledgments

Old Dominion University’s Climate Change and Sea Level Rise Initiative (CSCSLRI) and the Center for Coastal Physical Oceanography (CCPO) provided partial support and computational facilities for this study.

References
