

Intercomparison of the Gulf Stream in ocean reanalyses: 1993–2010

Lequan Chi*, Christopher L.P. Wolfe, Sultan Hameed

School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY 11794-5000, USA

ARTICLE INFO

Keywords:

North Atlantic Ocean
Gulf Stream
Ocean reanalysis
Intercomparison

ABSTRACT

In recent years, significant progress has been made in the development of high-resolution ocean reanalysis products. This paper compares aspects of the Gulf Stream (GS) from the Florida Straits to south of the Grand Banks—particularly Florida Strait transport, separation of the GS near Cape Hatteras, GS properties along the Oleander Line (from New Jersey to Bermuda), GS path, and the GS north wall positions—in 13 widely used global reanalysis products of various resolutions, including two unconstrained products. A large spread across reanalysis products is found. HYCOM and GLORYS2v4 stand out for their superior performance by most metrics. Some common biases are found in all discussed models; for example, the velocity structure of the GS near the Oleander Line is too symmetrical and the maximum velocity is too weak compared with observations. Less than half of the reanalysis products show significant correlations (at the 95% confidence level) with observations for the GS separation latitude at Cape Hatteras, the GS transport, and net transport across Oleander Line. The cross-stream velocity structure is further discussed by a theoretical model idealizing GS as a smoothed PV front.

1. Introduction

A complete quantification of ocean variables through their horizontal and vertical extents is needed to initialize and validate ocean circulation models and the oceanic component of climate models. Errors in specifying the state of the ocean are, therefore, a major source of uncertainty in climate modeling at this time.

Numerical modeling of the global ocean has seen significant progress in recent years. Large errors do, however, remain; many of these errors can be attributed to inadequate grid resolution, poor parametrization of key processes, errors in initialization, and propagation of numerical errors. In an effort to circumvent these shortcomings, reanalysis products have been developed in which the models are constrained to be consistent with available observations using data assimilation procedures. Reanalysis products are expected to provide more accurate information than unconstrained numerical models, with the caveat that significant errors are likely to remain because of inadequate coverage of the ocean by observation networks (Balmaseda et al., 2015; Toyoda et al., 2017). In addition, inconsistencies among different reanalyses also exist due to differences in resolutions, model physics, and data assimilation methods.

Reanalysis intercomparisons are well-established in the atmospheric science community, but have only recently become common in oceanography. For example, Balmaseda et al. (2015) presented an ocean reanalysis intercomparison focused on global features of the ocean, such as heat content, sea level, steric height, surface heat fluxes, mixed

layer depth, salinity, depth of 20 °C isotherm, and sea ice for the period 1993–2010. The authors compared several reanalyses with different resolutions, atmospheric fluxes or assimilation methods and found several areas of large uncertainty (differences among models) and systematic biases with respect to observations, pointing to the need for continuous model improvement. Through the comparison of several long-term reanalysis products, Stammer et al. (2010) found an increasing tendency of the spread across these products toward present. They suggested the convergence of different products in early years might be determined by boundary forcing and initial conditions, while the large discrepancies close to the end of the estimation period are due to the different assimilation approaches. Lee et al. (2009) reviewed existing reanalysis products and their application to meridional overturning circulation, pointing out the challenges in robustly estimating the uncertainties of models and observations. Stammer et al. (2016) reviewed applications of the ocean data assimilation in climate research, pointing out the necessity of fully coupled Earth system models for climate forecasts. The reader is referred to the special issue on reanalysis intercomparisons in volume 49, issue 3, of *Climate Dynamics* for a more comprehensive overview of the current state of ocean reanalysis intercomparisons.

The intercomparison presented in this paper is narrower in scope than the previous efforts, with a focus on the Gulf Stream (GS) from the Florida Straits to south of the Grand Banks. The GS is not only the western boundary current of the North Atlantic subtropical gyre, but also a branch of the Atlantic meridional overturning circulation

* Corresponding author.

E-mail address: lequan.chi@stonybrook.edu (L. Chi).

(AMOC). As such, the GS has a major influence on the dynamics and climate of the North Atlantic, as well as the climates of eastern North America and Europe (Kelly et al., 2010; Kwon et al., 2010; Palter 2015). A practical motivation for the present intercomparison is to provide an evaluation of reanalysis products for those involved in GS research.

Furthermore, the goal of ocean modeling—and reanalysis in particular—is to produce a model of reality. The underlying hope is that these models will eventually converge on the same result—in which case, we can say that we've learned something about the physical system. The fact that reanalyses often disagree with observations and each other tells us that we have some distance to go before realizing this goal. Therefore, a detailed validation of reanalysis products is valuable in its own right.

The GS is one of the most densely observed oceanic features in the world. One expects these observations should place a particularly heavy constraint on reanalysis products compared to more sparsely observed parts of the ocean and that this would tend to minimize inter-model differences. In this sense, the GS represents perhaps the most favorable setting in which to assess the skill of the reanalysis products—if they fail to adequately represent this observation-dense region of the ocean, there is little reason to expect that they will provide useful results in data-poor regions, such as the Southern Ocean.

Several previous works have compared the GS in numerical models but with different focus than this paper. Hurlburt and Hogan (2000) and Chassignet and Xu (2017) evaluated the GS in a series of numerical models with horizontal resolution varying from $\sim 1/10^\circ$ to $\sim 1/50^\circ$, concluding that the GS simulation benefits from higher resolution. Specifically, a significant improvement of the GS simulation is found once the resolution is sufficiently fine to resolve submesoscales ($\sim 1/50^\circ$) (Chassignet and Xu, 2017). Hurlburt et al. (2011) investigated the multi-year mean GS path and its separation near the Cape Hatteras in several eddy-resolving numerical models and found a very positive impact on the GS path simulation from data assimilation.

In this paper, we evaluate the mean state and variability of several GS features in the reanalysis products and observational data listed in Section 2. Specifically, the following aspects will be discussed in Section 3: the Florida Current transport (FCT); the separation of GS from the coast at Cape Hatteras; the GS transport and structure along the Oleander Line (a transect from New Jersey to Bermuda); and the positions of the GS path and the GS north wall. In Section 4, a simple theoretical model is employed to explain the cross-stream structure of the Gulf stream, and the overall performance of the numerical products is discussed. Finally, the paper is summarized in Section 5.

2. Datasets

2.1. Reanalyses

Table 1 lists the reanalysis datasets used in this study. All of these products are based on one of four ocean general circulation models (OGCMs):

- (a) MIT general circulation model (MITgcm, Marshall et al., 1997);
- (b) HYbrid Coordinate Ocean Model (HYCOM, Bleck 2002);
- (c) Modular Ocean Model (MOM, Pacanowski and Griffies 2000; Griffies et al., 2004);
- (d) Nucleus for European Modelling of the Ocean (NEMO, Madec 2008).

The products based on NEMO use the Louvain-la-Neuve Sea Ice Model (LIM) and those based on MOM4 and MOM5 use the Sea Ice Simulator (SIS) to simulate sea ice. Products based on HYCOM and the MITgcm use those OGCM's built-in sea ice models, while OFES and GODAS are run without sea ice. Further, OFES and ECCO2 are unconstrained; that is, they are run without assimilated data. They are included in this research as examples of free-run results.

The reanalysis products can be divided into three groups based on their horizontal resolution:

- (a) **eddy-resolving**, with resolutions of $\sim 0.1^\circ$ or finer (HYCOM and OFES),
- (b) **eddy-permitting**, with resolutions of $\sim 0.25^\circ$ (ORAP5, C-GLORSv5, GLORYS2v4, UR025.4, SODA3.3.1 and ECCO2),
- (c) and **coarse resolution**, with resolutions of $\sim 0.5^\circ$ or coarser (CFSR, ECCOv4r2, GODAS, GECCO2 and ECDav3.1).

The eddy-resolving and eddy-permitting products will be considered as **high-resolution numerical products** in this paper. An overlapping period of all reanalysis products, 1993–2010, is used in all following analysis unless otherwise noted. The numerical products are always listed in order of their resolution (from high-resolution to low-resolution).

2.2. Observations

The observationally based datasets used to verify the numerical models are described in this section.

2.2.1. EN4 (version 1.1)

EN4 (version 1.1) is a 1° gridded temperature and salinity dataset based on objective analysis of observations since 1900. It is produced by the UK Met Office Hadley Centre (Good et al., 2013) and widely used in oceanic and climate research (Armour et al., 2016; Dieng et al., 2015; Häkkinen et al., 2015). In this paper, it is considered as the reference—or “true”—temperature when in-situ observations are unavailable.

2.2.2. Absolute dynamic topography (ADT)

Absolute dynamic topography (ADT) is the sea surface height (SSH) with respect to the geoid. The filtered along-track ADT (product version 5.0) and the all-satellite-merged 0.25° gridded ADT (product version 5.0) are used to detect the GS path. Both of them are distributed by Aviso,¹ and the gridded ADT product is referred as **Aviso-ADT** in this paper.

2.2.3. Armor3D (v3.1)

Armor3D (v3.1) consists of monthly-mean gridded temperature, salinity, SSH, and horizontal velocity reconstructed from observations at 0.25° resolution in the upper 1500 m. The temperature below sea surface is based on synthetic analyses and optimal interpolation. First, a synthetic temperature anomaly field is derived from sea level anomaly (SLA) and sea surface temperature (SST) via multiple linear regression. Then, the final temperature field is generated by combining the synthetic results (temperature anomaly plus climate mean temperature) and in-situ observations through optimal interpolation (Guinehut et al., 2012). The SSH and horizontal velocity data (computed from the thermal wind equation) is borrowed from its companion dataset, **Surcouf3D** (Mulet et al., 2012). Please note that Armor3D is a gridded product derived from observations. It is included to complement the EN4 data set and provide an estimate of observational errors.

2.2.4. The Florida Current transport (FCT)

The Florida Current transport (FCT) is well observed via underwater cables and cable-calibration cruises since 1982. Daily FCT data is available from the project website.² One large gap during 1998–2000 and a few small gaps exist in the records. Meinen et al. (2010) reviewed the historical FCT observations, suggesting that an accurate estimate of

¹ The altimeter products were produced by Ssalto/Duacs and distributed by Aviso, with support from Cnes (<http://www.aviso.altimetry.fr/duacs/>).

² <http://www.aoml.noaa.gov/phod/floridacurrent/index.php>.

Table 1

Reanalysis products used in this paper. The third column gives the resolution and temporal coverage of the product. (Only the years with at least 6 months of data are counted). The last column of the table provides the major assimilation method used in the corresponding dataset and which variables are assimilated in the dataset, where T/S is vertical temperature/salinity profiles, SIC is sea ice concentration, and SIT is sea ice thickness.

Name	Forcing	Resolution and Period	Oceanic Model	Assimilation method and assimilated data
HYCOM Reanalysis expt-19 (Allard et al., 2013)	CFSR	0.08° (1993–2012)	HYCOM (Energy loan Sea ice)	3DVAR (SST/SLA/T/S)
OFES (Sasaki et al., 2008)	NCEP	0.1° (1950–2011)	MOM3 (No sea ice)	Unconstrained (N/A)
ORAP5 (Zuo et al., 2015)	ERA-Interm	0.25° (1979–2013)	NEMOv3.4 + LIM2	3DVAR (SST/SLA/SIC/T/S)
C-GLORSv5 (Storto and Masina 2016)	ERA-Interm	0.25° (1980–2015)	NEMOv3.2 + LIM2	3DVAR (SST/SIC/SIT ^a /SLA/T/S)
GLORYS2v4 (Garraic and Parent, 2017)	ERA-Interm	0.25° (1993–2015)	NEMOv3.1 + LIM2	Kalman Filter + 3DVAR (SST/SLA/SIC/T/S)
UR025.4 (Haines et al., 2013)	ERA-Interm	0.25° (1989–2010)	NEMOv3.2 + LIM2	OI (SST/SLA/SIC/T/S)
SODA 3.3.1 (Carton et al., 2018)	MERRA2	0.25° (1980–2015)	MOM5 + SIS	OI (Unknown)
ECCO2 (Menemenlis et al., 2008)	JRA-25	0.25° (1992–present)	MITgcm (/w sea ice)	Unconstrained ^b (N/A)
CFSR (Saha et al., 2010)	coupled model	0.5° (1979–2010)	MOM4 + SIS	3DVAR (SST/SIC/T/S)
ECCOv4r2 (Forget et al., 2015)	ERA-Interm	1° (1992–2011)	MITgcm (/w sea ice)	4DVAR (SST/SLA/T/S)
GODAS (Behringer and Xue, 2004)	NCEP Reanalysis 2	1° (1980–present)	MOM3 (No sea ice)	3DVAR (SST/T)
GECCO2 (Köhl 2015)	NCEP Reanalysis 1	1° (1948–2014)	MITgcm (/w sea ice)	4DVAR (SST/SLA/T/S)
ECDAv3.1 (Chang et al., 2013)	coupled	1° (1961–present)	MOMv4 + SIS	Ensemble Kalman Filter (SST/T/S)

^a Derived from the Pan-Arctic Ocean Modeling and Assimilation System.

^b Several 1–2-year-long constrained products are publically available, but the continuous 24-year-long product is unconstrained (Menemenlis, 2016).

monthly mean FCT requires at least 20 daily observations. To get a continuous time series of monthly mean FCT, the transport in the months with less than 20 observations are derived from another FCT dataset based on satellite altimetry (Goni et al., 2017).

2.2.5. The Oleander project

The Oleander Project has observed temperature and velocity via XBTs and a ship-mounted-ADCP along a transect between New Jersey and Bermuda (the Oleander Line, black dashed line in Fig. 1) regularly since 1992 (Flagg et al., 1998; Rossby et al., 2014; Rossby and Gottlieb 1998). Velocity data is available to 400 m prior to 2005 and to 800 m after that, with a one-year gap in 2008. The XBT observation only covers half of the Oleander Line (from New Jersey to the GS) before 2009.

In addition to in-situ data, the Oleander Project also provides estimates of the GS position along the Oleander Line and six-month running-mean transports across the Oleander Line.

3. Comparison between observations and models

3.1. Florida Current transport

The climate mean of the FCT is 32.1 ± 0.2 Sv (Meinen et al., 2010)³ with annually averaged values varying between 29.0 Sv and 34.6 Sv (Sanchez-Franks 2015). The FCT is a combination of transports from North Atlantic subtropical gyre and the upper limb of Atlantic meridional overturning circulation (AMOC). If all upper limb AMOC transport crossed 26°N via Florida Strait, 17.2 Sv (transport of AMOC at 26°N, McCarthy et al., 2015) of the FCT is from AMOC, and the rest (14.9 Sv) is from the subtropical gyre. By analyzing water properties, Schmitz and Richardson (1991) find that the near-surface (temperature > 24 °C) flow and near-bottom (temperature between 7 and 12 °C)

flow across the Florida Strait are primarily AMOC components originating from South Atlantic, while the waters found mid-depth (temperature between 12 and 24 °C) are mainly from the North Atlantic subtropical gyre. Xu et al. (2012, 2016) find similar results in their numerical simulations. Szuts and Meinen (2017) discuss the fate of the waters passing through the Florida Strait, finding that the near-bottom water, the intermediate water on the west side, and part of the intermediate water on the east side contribute to the AMOC. In contrast, the near-surface water and the remaining intermediate water on the east side return to the subtropical gyre. The interannual variability of FCT during 1982–1998 has been explained by wind stress curl near 27°N and the NAO via a Rossby wave mechanism; however, this mechanism fails to explain the variability of the FCT outside of this period (DiNezio et al., 2009; Meinen et al., 2010). Even though the Florida Current marks the upstream end of the GS, Sanchez-Franks (2015) shows that FCT is uncorrelated with GS transport across the Oleander Line.

Fig. 2 shows boxplots of annual mean FCT during 1993–2010 from direct observations via underwater cables, an observationally based dataset (Armor3D), and numerical models (in order of coarser resolution from left to right). The means and standard deviations of the FCT are also shown in Table 2, which summarizes the results of this study. All high-resolution models (horizontal resolution $\leq 0.25^\circ$) have lower median and mean FCT than the observations, while some coarse-resolution models obtain higher median and mean FCT values. However, the Interquartile Range (IQR) and median absolute deviation (MAD) of the FCT is much larger in the numerical models (> 0.99 and > 0.51, respectively) than in observations (0.66 and 0.29, respectively). The four models with mean FCT closest to the observations are ORAP5, CFSR, ECCOv4 and ECDAv3.1. The four models with standard deviation of FCT closest to the observations are HYCOM, OFES, GLORYS2v4 and ECCO2.

The annual mean FCTs in most numerical models are significantly correlated to the observations at 95% level with a highest coefficient of 0.74 (GLORYS2v4), while large spread exists between different seasons (Table 3). Generally, the numerical models perform better in summer

³ The ± 0.2 Sv is the statistical standard error of the mean.

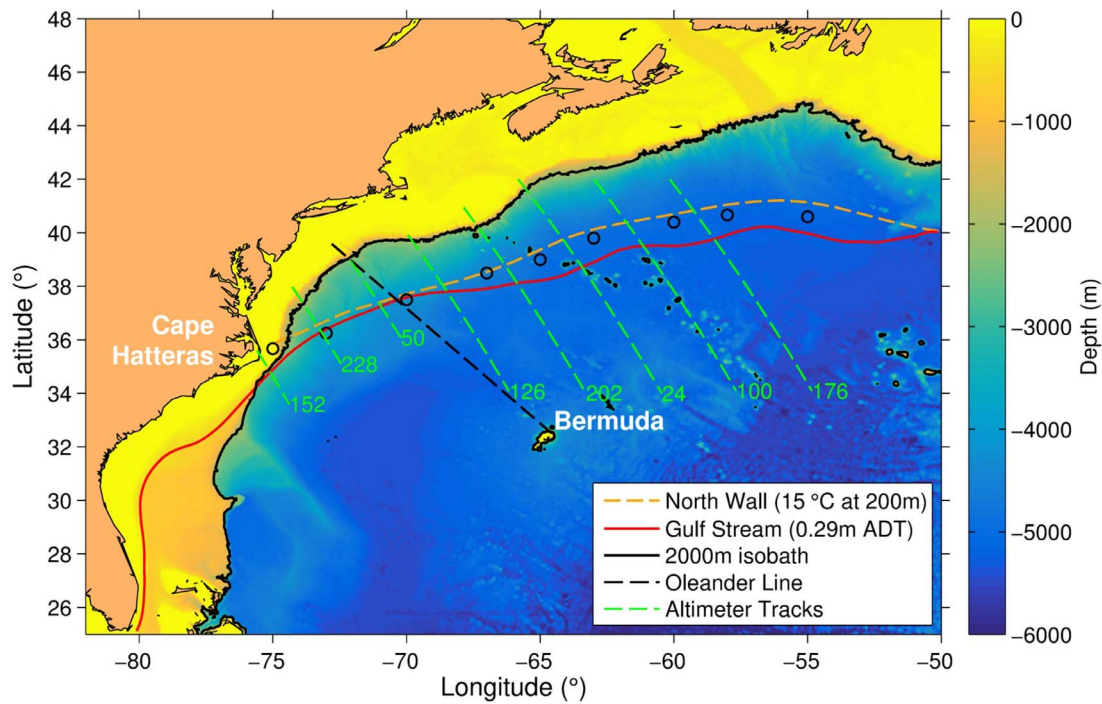


Fig. 1. The GS path based on 0.29-m ADT contour from satellite altimeter (solid red line) and the GS north wall derived from the 15 °C isotherm at 200 m from EN4 (dashed orange line). The black dashed line indicates the Oleander Line (a transect between New Jersey and Bermuda). The green dashed lines indicate tracks of satellite altimeter with track numbers noted at the south end of each track. The background color shows the topography and the black solid line is the 2000 m isobath. The circles mark the locations used in the calculation of Joyce's Gulf Stream index (Joyce et al., 2009). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

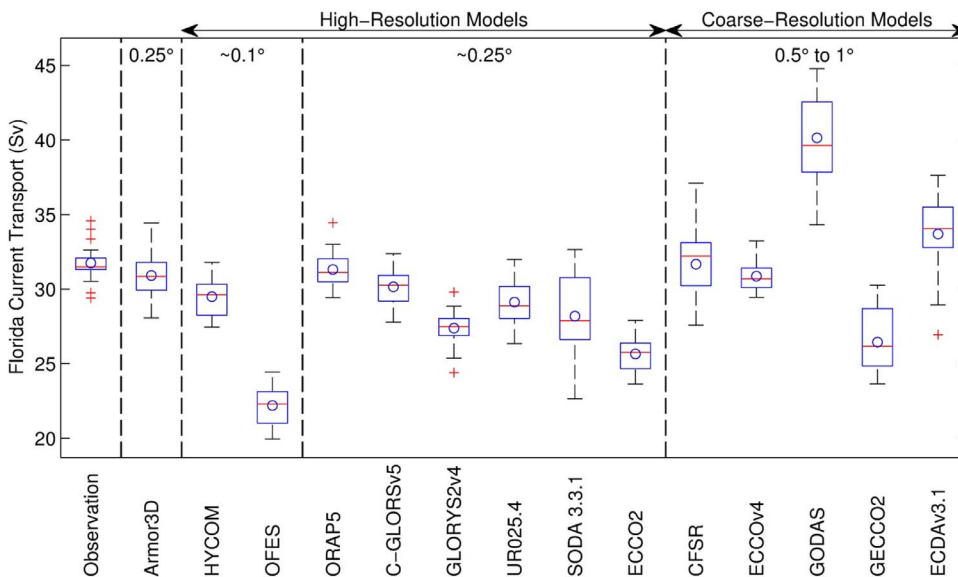


Fig. 2. Box plots of annual mean Florida Current transport from different sources. The horizontal lines in the boxes indicate medians, and the circles indicate mean values. The lower and upper boundaries of the boxes indicate Q_1 (the 25% percentile) and Q_3 (the 75% percentile), and values more than $1.5(Q_3 - Q_1)$ away from the box boundaries are marked by red crosses. The horizontal resolution of the numerical models gets coarser from left to right and is marked on the top of the figure.

(JAS) and autumn (OND) than in winter (JFM) and spring (AMJ). Only 1 (0) of the 13 numerical models shows significant correlation with the observations at zero lag in winter (spring), while in summer (autumn), 6 (9) models show significant correlations. At 90% confidence level, GLORYS2v4 is the only reanalysis product in which the FCT is significantly correlated to the observations in all four seasons. A similar seasonal pattern also exists in the observational dataset Armor3D. The seasonal dependence of correlation indicates that the mechanism driving FCT variability may change with season.

3.2. The Gulf Stream separation at Cape Hatteras

After the Florida Straits, the GS flows northward as a narrow, stable

current along the coast until Cape Hatteras, where it turns eastward and separates from the coast. This separation point is considered a key test of GS simulations as it is not reproduced well by many models. Hurlburt and Hogan (2000) suggest that $1/32^\circ$ horizontal resolution is required for the robust simulation of the separation in unconstrained models. Following Gangopadhyay et al. (1992), the separation latitude (Fig. 3) used in this paper is defined as the latitude where GS crosses the 2000-m isobath.⁴ The mean separation latitude during 1993–2010 is 34.25°N with a standard deviation of 0.15° . This is $\sim 1^\circ$ south of the separation latitude found by Gangopadhyay et al. (1992). However, the GS path used in their research is detected via frontal analysis from

⁴ The details of GS path detection are introduced in Section 3.4.1.

Table 2

Summary of the GS properties in different products. All results are derived from annual mean values during 1993–2010 unless otherwise noted. The bold numbers indicate the reanalysis product closest to observations; the four reanalyses closest to the observations at each aspect are marked by shaded backgrounds. The value for Armor3D is also marked by shading background if its performance is comparable with the four best reanalyses in the corresponding aspect. For correlation coefficients, only those significant at 95% level are listed.

		Observations	Armor3D	HYCOM	OFES	ORAP5	C-GLORSv5	GLORYS2v4	UR025.4	SODA 3.3.1	ECCO2	CFSR	ECCOv4r2	GODAS	GECCO2	ECDav3.1
Florida Current transport	r	-	0.63	0.63	-	0.49	0.55	0.74	-	-	-	0.57	0.61	-	-	-
	mean (Sv)	31.9	31.1	29.5	22.0	31.2	30.3	27.6	29.2	28.4	25.4	31.6	30.7	39.7	26.4	33.3
	Std (Sv)	1.07	1.51	1.16	1.18	0.92	1.26	1.03	1.59	2.60	1.09	2.20	0.87	2.97	2.00	2.83
Separation latitude	Event 2010 ^a	4.7	4.4	4.1	-	-	-	-	3.4	-	-	-	-	-	-	-
	Monthly RMSE ^b	-	0.05	0.11	0.29	0.14	0.19	0.12	0.12	0.12	0.24	0.12	0.16	0.13	0.09	0.12
	Mean (°N)	34.25	34.31	34.08	33.94	34.30	34.05	34.30	34.17	34.02	33.33	34.34	33.80	36.55	36.83	33.18
	Std (°)	0.15	0.18	0.20	0.20	0.41	0.09	0.11	0.15	0.25	0.37	0.40	0.08	0.30	0.53	0.11
	r	-	0.98	0.72	-	-	-	0.74	-	-	-	-	-	-	-	-
GS path	AEA-GS (10 ¹⁰ m ²)	-	1.0	5.4	28.9	16.1	13.8	6.4	7.9	14.8	23.2	14.2	9.7	14.9	17.5	10.3
GS North Wall	AEA-NW (10 ¹⁰ m ²)	-	9.75	11.02	19.20	17.83	11.40	10.62	9.93	12.84	19.78	16.18	7.51	13.06	17.99	10.63
	r ^c	-	0.87	0.79	-	-	0.83	0.64	0.80	0.66	-	0.79	0.65	0.81	-	0.62
Oleander Line	Horizontal and vertical structure ^d	Max velocity (m s ⁻¹)	1.91	1.10	1.08	0.86	0.87	0.73	1.02	0.90	0.81	0.60	-	-	-	-
		GS Axis bias (km) ^e	-	-15.0	-2.5	-3.5	3.0	-3.0	-4.0	-3.5	2.5	-2.0	-	-	-	-
		Mean kurtosis ^f	3.16	2.79	2.62	2.53	2.83	2.46	2.67	2.50	2.55	2.37	-	-	-	-
		Mean skewness ^g	0.58	0.26	0.18	-0.08	0.20	0.08	0.20	0.05	0.09	0.11	-	-	-	-
	15°C isotherm	Position bias (km) ^h	-	14.0	6.4	175.9	208.1	46.6	7.3	11.6	26.2	277.2	89.2	105.4	132.8	91.5
	GS transport	Mean (1 × 10 ⁴ m ² s ⁻¹)	13.59	12.58	12.70	10.97	9.45	9.89	12.27	12.67	11.88	7.83	7.20	8.73	7.83	7.02
		Std (1 × 10 ⁴ m ² s ⁻¹)	0.78	0.91	1.15	1.21	0.77	1.00	1.20	1.27	1.75	1.50	0.63	0.31	1.78	1.61
		r	-	0.56	0.55	-	-	-	0.59	-	-	-	-	-	0.49	0.40
	Net transport	Mean (1 × 10 ⁴ m ² s ⁻¹)	8.37	7.97	8.25	7.93	7.56	7.65	8.40	8.56	8.27	7.34	6.78	7.96	6.86	5.85
		Std (1 × 10 ⁴ m ² s ⁻¹)	1.08	0.74	0.88	1.18	0.52	0.53	0.70	0.69	0.70	0.97	0.74	0.29	1.82	1.64
		r	-	0.34	0.41	-	-	-	0.38	-	-	-	-0.36	-	0.36	-
	GS latitude	Mean (°N)	37.57	37.45	37.43	36.39	38.40	37.25	37.45	37.45	37.24	36.37	37.76	37.69	37.91	37.18
		Std (°)	0.18	0.19	0.21	0.48	0.39	0.25	0.22	0.25	0.31	0.52	0.43	0.35	0.73	0.41
		r	-	0.90	0.88	-	-	-	0.68	0.72	-	-0.52	0.49	0.67	-	-

^aThe values are the southernmost monthly latitudes in 2010 compared to their mean value in units of their standard deviation. Only values exceeding 3σ are listed.

^bThe RMSE is derived from the 1993–2010 mean separation latitude in each month.

^cThe correlation between RGS-E8 and GSI-E8 during 1993–2010.

^dThe structure is discussed during 2005–2010 due to the availability of in-situ observations.

^eThe averaged bias of the mean maximum velocity positions in reanalysis products at 55 m, 200 m, 400 m, 600 m. Negative values indicate northward biases.

^{f,g}Averaged kurtosis (skewness) of the horizontal GS velocity structure at 55 m, 200 m, 400 m and 600 m; results from coarse-resolution reanalyses are omitted.

^hAveraged at 200 m, 400 m and 600 m during July 2009–December 2010.

satellite infrared images, which is closer to the definition of GS north wall (further discussed in Section 3.4.2). Since the GS north wall is always located north to the GS path, the difference between the separation latitude derived in this paper and derived by Gangopadhyay et al. (1992) is probably a result of the different definitions of the GS path.

The observed separation latitude via Aviso-ADT data is highly stable except for a slight northward movement during 1998–2000 and an extreme southward shift in 2010 (Fig. 3). This extreme event (Event, 2010) lasted about 3 months with minimum monthly latitude of 32.2°N in February 2010, which is 4.7σ south of its mean position. The cause of this event is still unknown, but it might be related to the sharp decrease in AMOC transport and the extremely low NAO event in the winter of 2009–2010 (McCarthy et al., 2012; Bryden et al., 2014). The DJF-mean NAO index in 2010 was -1.67, which is the lowest value since 1950. Only Armor3D and 3 of the 13 numerical models, HYCOM, GLORYS2v4 and UR025.4, can reproduce this event with more than a 3σ southward shift from its mean latitude.

Fig. 4 shows the climate mean separation latitudes for each month with boxplots derived from Aviso-ADT. An annual cycle with a maximum latitude in September is clearly visible. Although most of the numerical

models can reproduce a seasonal cycle with low separation latitude in winter and spring, and high separation latitude in summer and autumn, only a few of them can reproduce the sharp peak in September. To evaluate the seasonal cycle in each numerical model, root-mean-square error (RMSE) is calculated for each dataset (Table 2) as

$$RMSE = \frac{1}{12} \cdot \sqrt{\sum_{i=1}^{12} (Slat_{norm,model,i} - Slat_{norm,obs,i})^2},$$

where $Slat_{norm,obs,i}$ is the 1993–2010 mean observed separation latitude for month i and the subscript *model* indicates that the separation latitude is from a numerical model. The separation latitudes were normalized before the calculation of RMSE. Lower RMSE indicates a seasonal cycle closer to the observed one. The four numerical models with smallest RMSE are HYCOM (0.11°), GLORYS2v4 (0.12°), UR025.4 (0.12°), SODA 3.3.1 (0.12°), CFSR (0.12°), GECCO2 (0.09°) and ECDav3.1 (0.12°).⁵

To evaluate the interannual variability of the numerical models, boxplots of annual mean separation latitudes from observations and all

⁵ Seven models are listed here because five of them share similar value of RMSE (0.12°).

Table 3

Correlation coefficients of FCT from observations and other datasets during 1993–2010. The values significant at 95% confidence level are in bold with a shaded background.

Dataset	Annual	Winter (JFM)	Spring (AMJ)	Summer (JAS)	Autumn (OND)
Armor3D	0.63	0.43	0.34	0.52	0.71
HYCOM	0.63	0.33	0.12	0.52	0.77
OFES	0.28	0.01	-0.28	0.38	0.38
ORAP5	0.49	0.45	0.11	0.25	0.55
C-GLORSv5	0.55	0.58	0.13	0.45	0.79
GLORYS2v4	0.74	0.44	0.44	0.62	0.71
UR025.4	0.22	-0.02	0.22	0.70	0.31
SODA 3.3.1	-0.27	-0.12	-0.03	0.38	0.02
ECCO2	0.10	-0.09	-0.11	0.20	0.42
CFSR	0.57	0.18	0.33	0.35	0.55
ECCOv4	0.61	0.42	0.09	0.29	0.72
GODAS	0.47	0.27	0.06	0.61	0.56
GECCO2	0.41	-0.05	0.09	0.57	0.53
ODA	-0.06	0.29	-0.17	0.13	-0.03

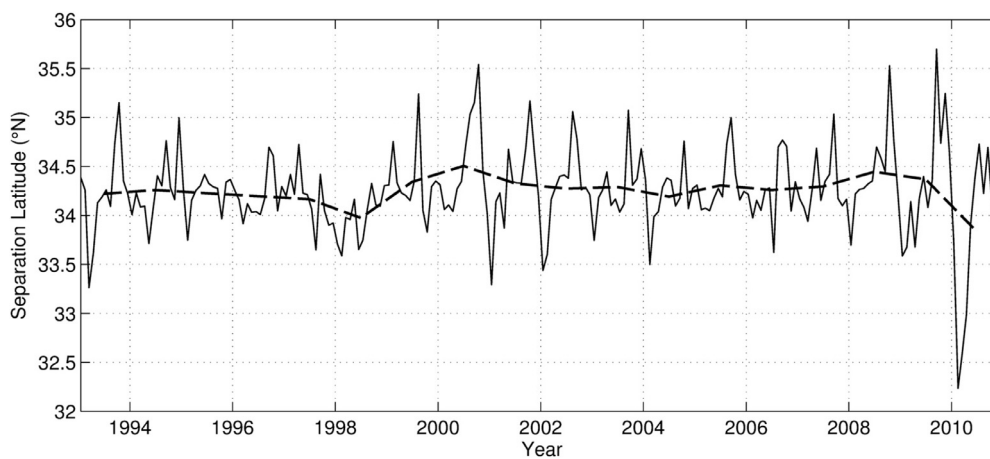


Fig. 3. Monthly (thin line) and annual (thick line) GS separation latitude from Aviso $\frac{1}{4}^\circ$ gridded ADT data.

datasets are shown in Fig. 5. The GS in GODAS and GECCO2 does not separate from the coast at Cape Hatteras, while in ECCO2, ECCOv4 and ECDAv3.1, the GS separates from the coast too early. The four 1° -resolution models (ECCOv4, GODAS, GECCO2 and ECDAv3.1) have the largest mean or median bias except ECCOv4, consistent with the common agreement that 1° horizontal resolution is not sufficient for proper simulation of GS separation. (Chassignet and Marshall, 2008)

As expected, Armor3D is correlated with observations at better than 95% confidence, while the correlation coefficients from only 2 of the 13 numerical models—HYCOM and GLORYS2v4—are also significant at 95% confidence level.

3.3. Oleander Line

The Oleander Line (the black dashed line in Fig. 1) is a section between New Jersey and Bermuda with regular observations since 1992 using XBT and ship-mounted-ADCP⁶ (Flagg et al., 1998; Rossby et al., 2014). It is nearly perpendicular to the mean GS path, making the cross-

Oleander Line velocity a good approximation of the downstream velocity of the GS. The long-term continuous observations are very useful for comparison with the GS in numerical models. Note that the Oleander Line crosses the GS near a point where the latitudinal variability of the GS is minimal (Cornillon 1986). For a specific numerical model, its capacity to simulate this variability minimum will likely affect their agreement with Oleander observations.

3.3.1. GS vertical and horizontal structure

The structure of GS velocity across the Oleander Line is shown in a natural coordinate where maximum velocity is located at the origin (Figs. 6 and 7). The 75 kHz ADCP data—which can reach more than 800 m depth—is available since 2005. Only 46 cruises provide sufficient coverage to calculate the velocity distribution in the GS. These cruises are seasonally biased, with 31 of the cruises in May–September during 2005–2010. To compensate for this seasonal bias, only velocities from the corresponding months from Aromr3D and numerical models are used in this section.

The observed GS is about 200 km wide (distance between the first zero downstream velocity position north/south to the GS path at 55 m depth) at the Oleander Line. In high-resolution models, the GS is about

⁶ Data available at <http://www.po.gso.uri.edu/rafos/research/ole/index.html>.

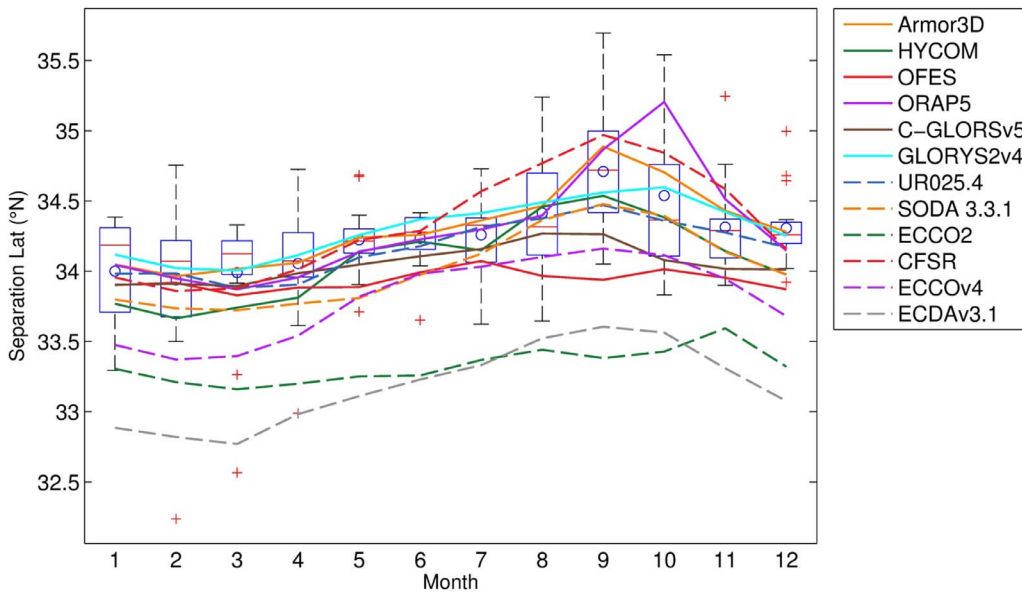


Fig. 4. Climate mean GS separation latitude for each month from observed ADT (Box plots). The boxplots are as in Fig. 2. Other solid and dashed lines indicate climate mean value from different datasets. The latitudes from GODAS (36.6°N) and GECCO2 (38°N) are too high to be shown in this figure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article)

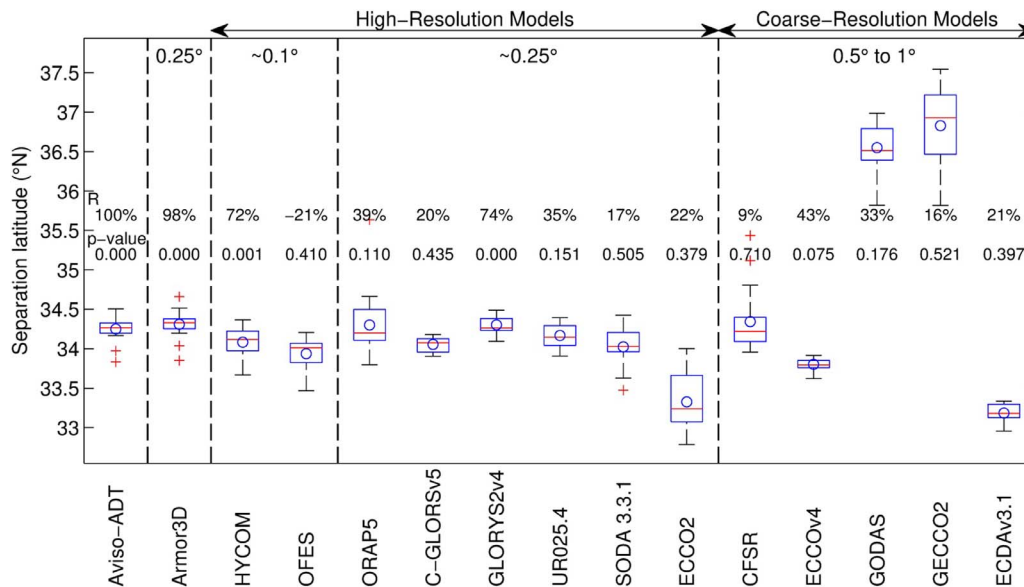


Fig. 5. Boxplot of annual mean GS separation following the convention of Fig. 2. The numbers on the figure are (upper) correlation coefficients and (lower) corresponding p-values to the observed separation latitude.

250–300 km wide, while it is much broader ($\gg 300$ km) in the coarse-resolution models. This is consistent with the extremely small downstream velocities in the coarse resolution models (maximum velocity across the Oleander Line is less than 0.3 m s^{-1}) and the correct order of net northward transport in the western boundary. Because the horizontal GS velocity structures in coarse-resolution models are very unrealistic, only the structures in high-resolution models will be discussed further.

The observed downstream velocity shows significant horizontal asymmetry (black lines in Fig. 6a–d). The velocity north of the GS path (the negative distance region in Fig. 6) decreases much faster than to its south (the positive distance region). To further discuss the horizontal structure quantitatively, kurtosis and skewness are employed by taking the downstream horizontal velocity as a probability distribution function. Negative velocity is ignored during this calculation. The asymmetry of the horizontal velocity decreases with depth, with a skewness that decreases from 0.88 at 55 m to 0.33 at 600 m. The sharp decrease in the velocity to the north of the maximum likely results from the

strong potential vorticity front near the north wall of the GS (see Section 4). The horizontal downstream velocity structure in both the observationally based Armor3D and most of the numerical models are nearly symmetric (Figs. 6 and 7). This is also shown by the average skewness of horizontal velocity structure at 55 m, 200 m, 400 m and 600 m depth (Table 2). The average skewness from observations is 0.58, while the largest skewness from numerical models is 0.20 (ORAP5 and GLORYS2v4).

Besides, the velocity decreases more slowly away from the peak in Armor3D and the numerical models compared to observations above 600 m. The kurtosis of the observed horizontal GS velocity structure (3.16) is larger than that in the numerical models (no more than 2.83); that is, the horizontal velocity distribution is compared to its tails in observations than in numerical models. In observations, the kurtosis of horizontal GS velocity decreases with depth. The ratio of observed kurtosis at 600 m and 55 m is 0.75. This vertically decreasing pattern is well captured by ORAP5, the ratio in which is 0.77. The ratio is 0.89 in Armor3D, 0.83 in GLORYS2v4 and no less than 0.90 in the other high-

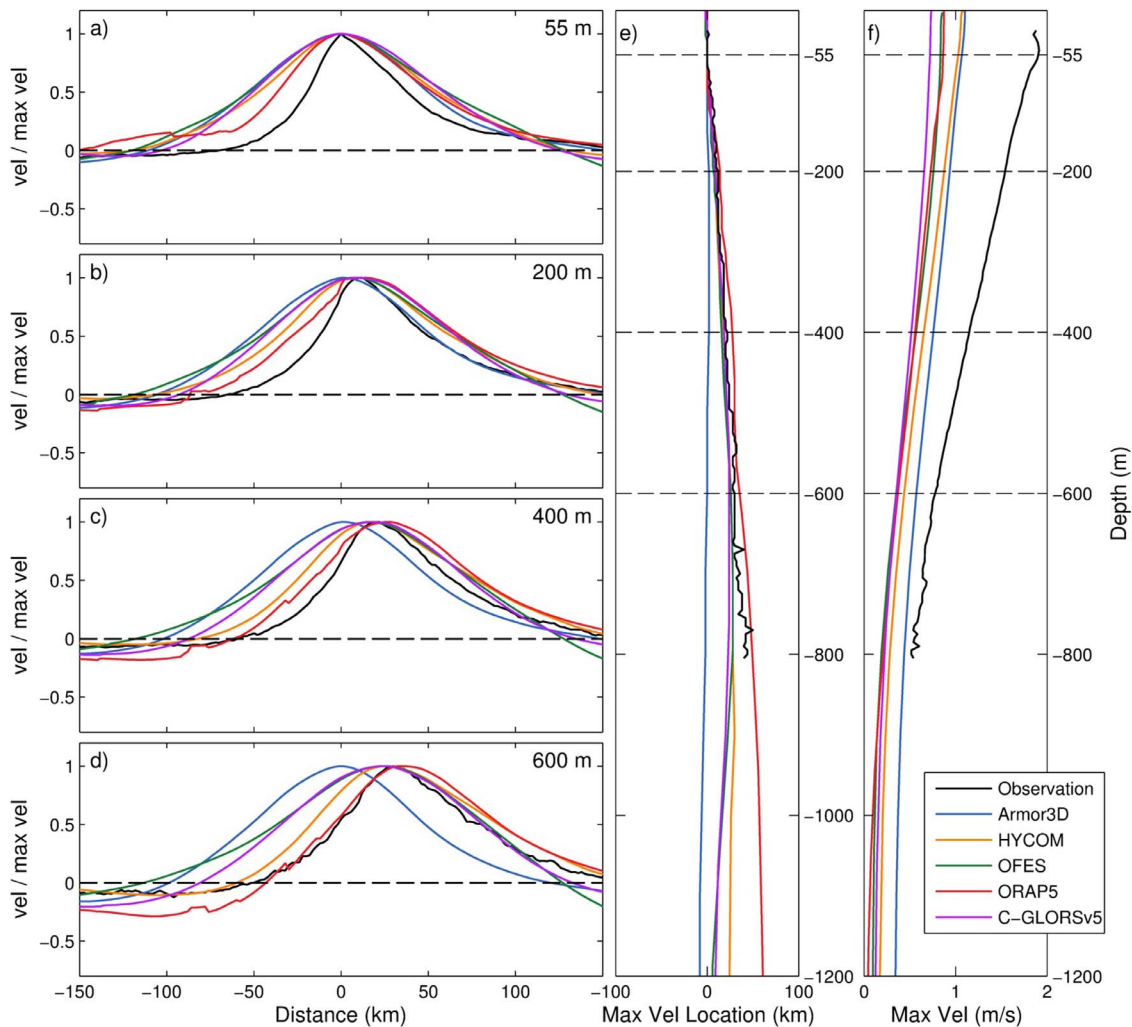


Fig. 6. 2005–2010 mean GS velocity in m s^{-1} across the Oleander Line in natural coordinates, where 0 km is the position of the maximum velocity at 55 m. The negative (positive) region is north (south) of the maximum velocity. Only data in months with in-situ observations are averaged. (a)–(d): velocity at different depth divided by its maximum value; (e) positions of the maximum velocity with depth; (f) the maximum velocity with depth.

resolution models.

A southward shift of the maximum velocity axis of the GS with depth (black line in Fig. 6e) can be found in ADCP observations. Ratsimandresy and Pelegri (2005) attributed this to the horizontal temperature gradient and the thermal wind relation. The GS axis shift with depth in most of high-resolution models is near the observed shift (Figs. 6e and 7e). A slightly northward shift of the GS axis is observed below 800 m in the high-resolution models, except ORAP5, in which a northward shift occurs below 1000 m (not shown); however, it is difficult to determine whether this is a real effect because this region is not covered by observations. The GS axis has a near constant position at all depths in Armor3D; this is probably due to the synthetic analysis method used in Armor3D (see Section 2 for more details).

Surprisingly, both the observationally based dataset Armor3D and the numerical models, including the eddy-resolving models, significantly underestimate the maximum downstream velocity across Oleander Line (Figs. 6f and 7f). The maximum downstream velocity is 1.91 m s^{-1} in the observations, while it is only 1.10 m s^{-1} in Armor3D, 1.08 m s^{-1} in HYCOM, 1.02 m s^{-1} in GLORYS2v4 and less than 1 m s^{-1} in the others. This bias is not significantly smaller in the eddy-resolving models (HYCOM and OFES) than in the eddy-permitting models. The underestimation of the peak velocity due to two factors: first, the numerical models underestimate the transport of the GS (Section 3.3.2); second, the GS is broader in the numerical models than in observations, so the already weak transport is spread over a larger

cross-sectional area.

To evaluate the temperature structure in the numerical models, Fig. 8 shows the averaged 15 °C isotherm in XBT observations, observationally based dataset Armor3D and numerical models along the Oleander Line during July 2009–December 2010. The 15 °C isotherm is shown because its position at 200 m depth is used as the position of the GS North Wall (Fuglister and Voorhis 1965; Joyce et al., 2000; Joyce et al., 2009). A sharp shoaling of the 15 °C isotherm can be found between 36.5°N and 38°N in the XBT observations (black solid line). Southeast of the sharp slope, the isotherm is horizontal for $\sim 1^\circ$ (35.5°N–36.5°N), then gently shoals toward the southeast. The position of the 15 °C isotherm along the Oleander Line is consistent with observations in most of the constrained high-resolution models, except in ORAP5 in which the 15 °C isotherm is several hundred meters below the observed one north of 37°N. The large bias in ORAP5 is due to the overshooting of the GS after Cape Hatteras. The isotherms in the two unconstrained products (OFES and ECCO2) at any specific time show a similar pattern as the observation: a sharp slope near the GS axis (not shown). The near-constant slope of the long-term mean isotherm in the two models is a result of their GS paths having more variability than observed. The slopes of isotherms in EN4, Armor3D and the numerical models are flatter with a shallower maximum depth than in the XBT observations. The slope in HYCOM is closest to the observations, and the slopes from Armor3D, GLORYS2v4, UR025.4 and SODA 3.3.1 are closer to the XBT observations than that from EN4. As with the down-

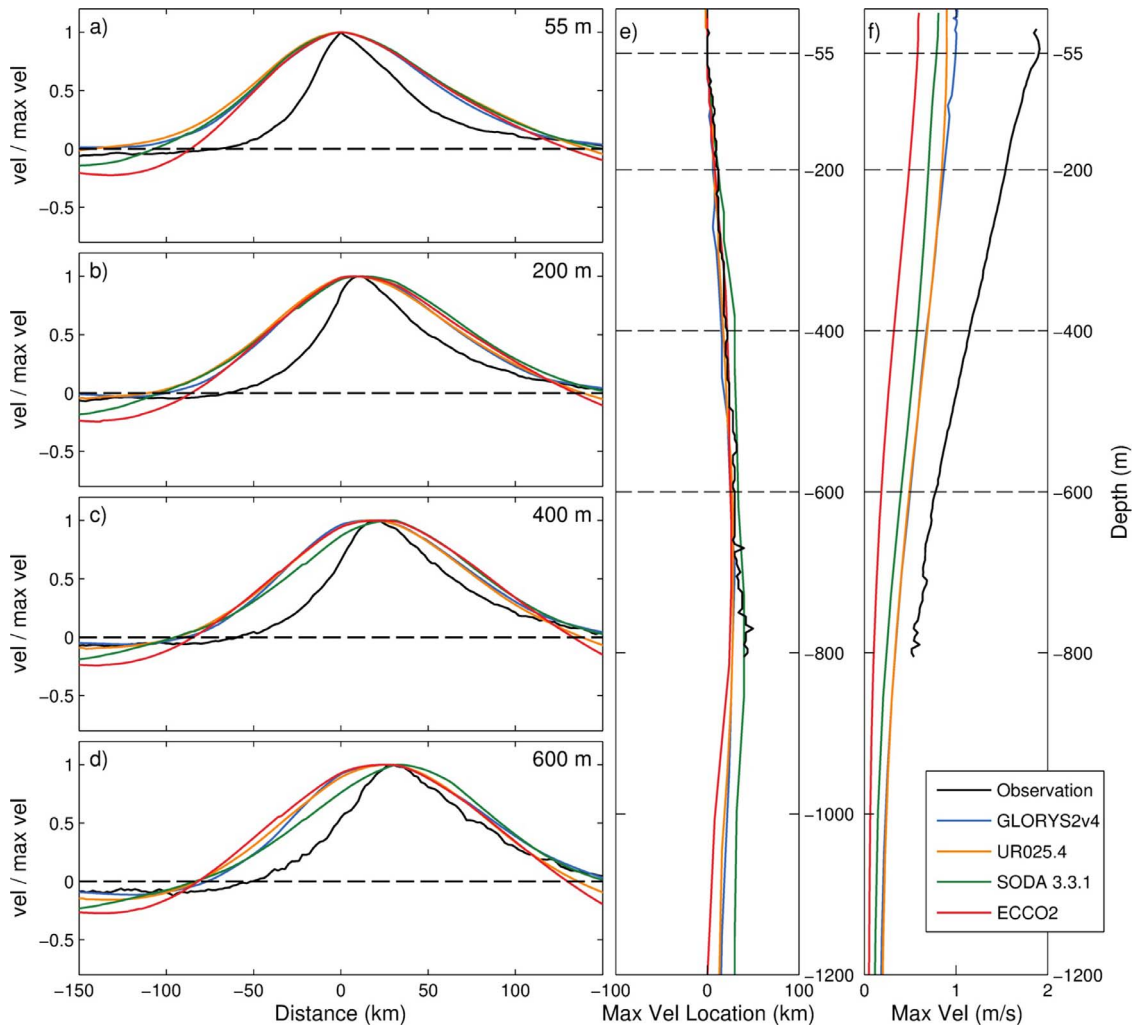


Fig. 7. As with Fig. 6, but for different models.

stream velocity along the Oleander Line, the isotherms in all coarse-resolution models are flatted compared to the observations.

To compare the slope and its position in different products quantitatively, the horizontal absolute distance between the 15 °C isotherms in the observationally based *Armor3D*, numerical models and in observations along the Oleander Line is calculated at 300 m, 400 m and 500 m. The distance averaged over the three layers is shown in Table 2. The averaged distance indicates a combination bias of both the mean position and the slope of the isotherm in each dataset. The 15 °C isotherm in *HYCOM* is closest to the observation with only a 6.4 km bias. The averaged distance for *EN4* is 62.6 km, which is larger than most high-resolution models and observationally based dataset *Armor3D*.

3.3.2. Gulf Stream transport across the Oleander Line

Transport across the Oleander Line consists of three parts: the Slope Sea Current (part of the Northern Recirculation Gyre), the GS, and the Southern Recirculation Gyre in the Sargasso Sea. *Rosby et al. (2014)* estimated GS transport across the Oleander Line by the velocity at ~55 m, which is just below the Ekman layer. Based on previous observations, they estimated that the GS transport integrated to 2000 m was 700 times the transport at 55 m. To be consistent with the observed transports, the transports across the Oleander Line in numerical models are also estimated from the velocity at a single layer at 55 m (unit: $\text{m}^3 \text{s}^{-1} \text{m}^{-1}$, or $\text{m}^2 \text{s}^{-1}$). Thus, the “GS transport” is the velocity at 55 m normal to the Oleander Line, integrated between where the normal

velocity first reaches zero from either side of the GS axis. The Slope Sea transport is the normal velocity at 55 m integrated from the northern edge of the GS to continental shelf (39.31°N, 72.57°W) and the Sargasso Sea transport is the normal velocity at 55 m integrated from the southern edge of the GS to near Bermuda (32.78°N, 64.89°W). The “net transport” is the sum of the transport in Slope Sea, the GS and the Sargasso Sea. Note that both the Slope Sea Current and the current in the Sargasso Sea are southwestward, and a large part of the two southwestward currents will join the GS. As a result, the net transport across the Oleander Line is smaller than the GS transport. Transports from observationally based *Armor3D* and numerical models are smoothed using a six-month running mean to be consistent with how the observational transports from the Oleander Project are processed. The transports from in-situ observations, *Armor3D* and numerical models are shown in Fig. 9; the errors of the transports in *Armor3D* and models are shown in Fig. 10.

All numerical models evaluated in this paper underestimate the GS transport, the Slope Sea transport and the Sargasso Sea transport across the Oleander Line (Fig. 9).⁷ The Slope Sea Current is missing in coarse-resolution models. And among all high-resolution models, only *HYCOM* and *UR025.4* reproduce the Slope Sea transport with the correct order of magnitude. The magnitude of Sargasso Sea transport is close to the observations in all high-resolution models except *ECCO2*.

⁷ In fact, the mean and median GS transports of the numerical models are smaller than more 75% of the observed transport values.

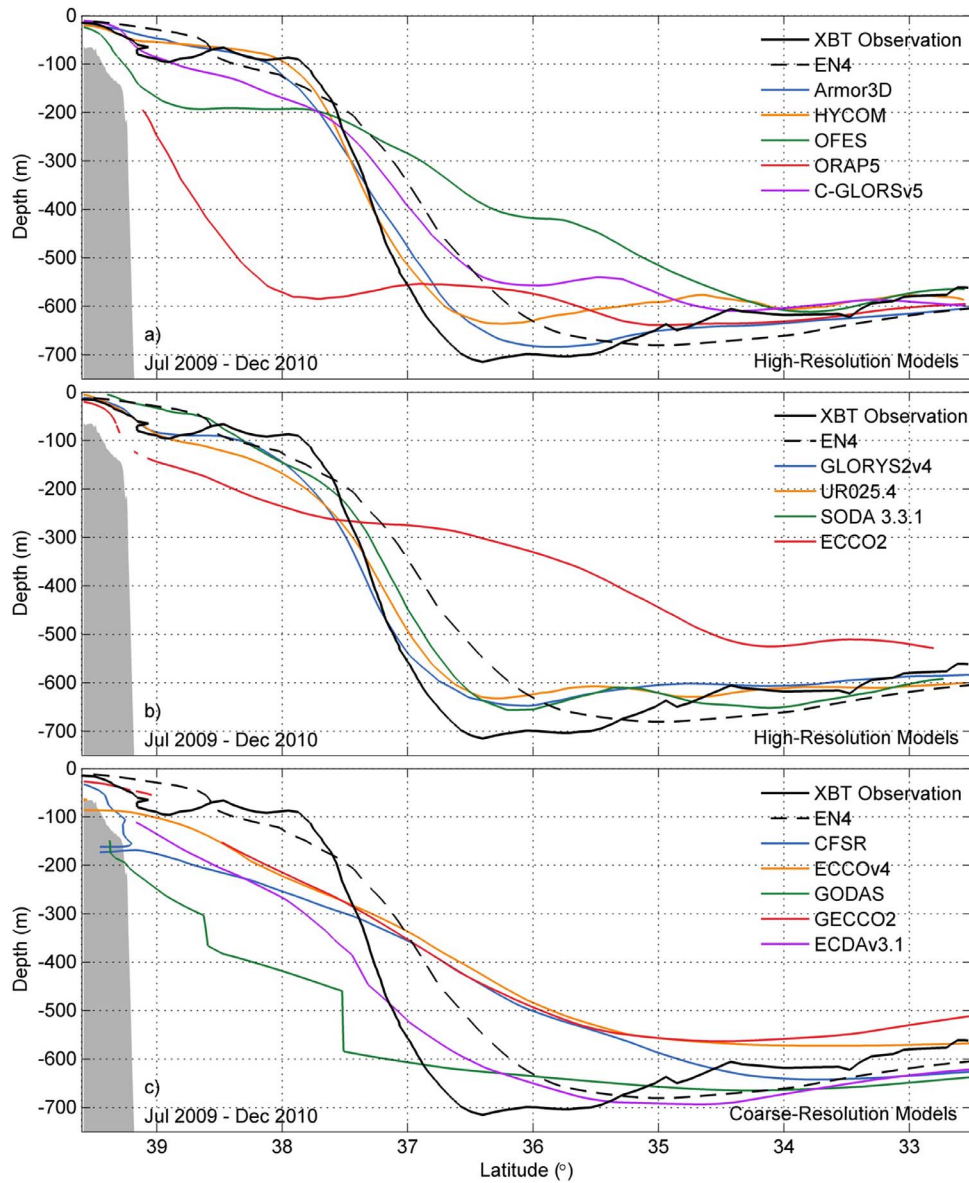


Fig. 8. Mean 15 °C isotherm along Oleander Line during July 2009 and December 2010.

The biases of the net transport in the numerical models are much smaller than the GS transport, which indicates that the GS transport bias in numerical models is mainly due to the underestimated transports of the two southwestward currents. More specifically, the error of the Slope Sea transport contributes more to the GS transport bias than the error of Sargasso Sea transport in most high-resolution models. In contrast, the error of the Sargasso Sea transport contributes more to the GS transport bias than the error of Slope Sea transport in the coarse-resolution models, even though the Slope Sea Current is missing in these models. (Fig. 10).

Similar results are shown by the time series of GS transports and net transports (Figs. 11 and 12). Generally, the GS transports in the high-resolution models are between the observed GS and net transports, while both transports in the coarse-resolution models are closer to the observed net transport. The observed net transport is locally maximum around January 2005, while all numerical models—as well as the observationally based Armor3D—show decreasing net transport at this time. The reason for this discrepancy is unclear.

Four of the thirteen numerical models are correlated to the observations in GS transport at 95% confidence level: HYCOM (55%), GLORYS2v4 (59%), GECCO2 (49%), and ECDAv3.1 (40%). HYCOM

(41.0%), GLORYS2V4 (38%), CFSR (−36%), and GODAS (36%) are correlated to the observations in net transport across the Oleander Line at 95% confidence level. Note that CFSR's correlation is *negative*. Further, GLORYS2v4 is more highly correlated with the Oleander observations of both GS and net transport than the observationally based Armor3D.

3.3.3. GS position at Oleander Line

The GS position at the Oleander Line is defined as the position of maximum downstream velocity at 55 m (Sanchez-Franks 2015). However, some numerical models occasionally give two velocity maxima within a single monthly mean at the Oleander Line. This may be due to switching of the path of GS within the month. In such cases, the peak closest to the GS path detected via Aviso-ADT (see Section 3.4.1) is considered as the primary axis of the GS.

The annual mean GS latitude along the Oleander Line from in-situ observations, Armor3D, and other numerical models is shown in Fig. 13. The observed position does not have a trend. The GS latitude is not significantly biased in the numerical models except in OFES (unconstrained), ORAP5, ECCO2 (unconstrained), GODAS (after 2005), and GECCO2 (after 2002); it is, however, more variable in the

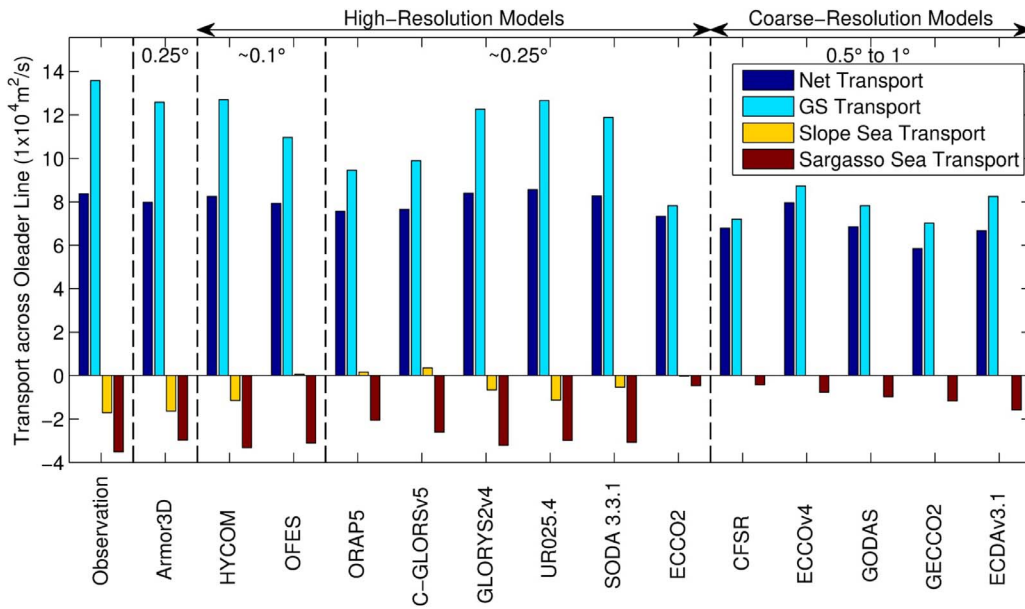


Fig. 9. The net transport, Gulf Stream transport, Slope Sea transport and Sargasso Sea transport across the Oleander Line. Negative values indicate southwestward transport.

numerical models than in the observations. The latitude fluctuations in 4 of the 13 models are significantly correlated with observations at the 95% level: HYCOM (88%), GLORYS2v4 (68%), UR025.4 (72%) and ECCOv4 (67%).

The seasonal cycle of the GS position along the Oleander line is not shown because it is weak in observations, with a small southward shift in Spring.

3.4. GS path and its North Wall

3.4.1. GS path

The “GS path” is the maximum velocity axis of GS. It is difficult to identify precisely and several different proxies are used in practice. One such proxy is the ADT contour of a specific value, which is usually 0.25 m (Rossby et al., 2014; Andres 2016). The 0.5-m sea surface height anomaly is also used in some cases (Gangopadhyay et al., 2016). However, the SSH reported by the numerical models is not generally

referenced to the same geoid used by observational datasets and so it is necessary to determine an appropriate SSH contour separately for each model. In each model, we choose the SSH contour that most frequently coincides with the location of maximum velocity at the 8 satellite-altimeter tracks shown in Fig. 1. More precisely, the chosen SSH contour is south (or north) of the velocity maximum at 50% of the space-time points defined by the altimeter tracks. When applied to the altimeter observations, this method gives the 0.29 m ADT contour as the GS path. This ADT value is sufficiently close to that used by Rossby et al. (2014) and Andres (2016) that we consider the method to be validated. In the following, we therefore use the 0.29-meter ADT contour to define the GS path in observations and similarly obtained SSH contours (ranging from −0.19 m to 0.31 m) to define the GS path in the numerical models.

The mean GS path from AvisoADT during 1993–2010 is shown in Fig. 14 (dashed black lines). The mean path of the GS is flat west of 65°W, with a few small undulations appearing to the east. In all constrained high-resolution models except ORAP5, the mean GS path

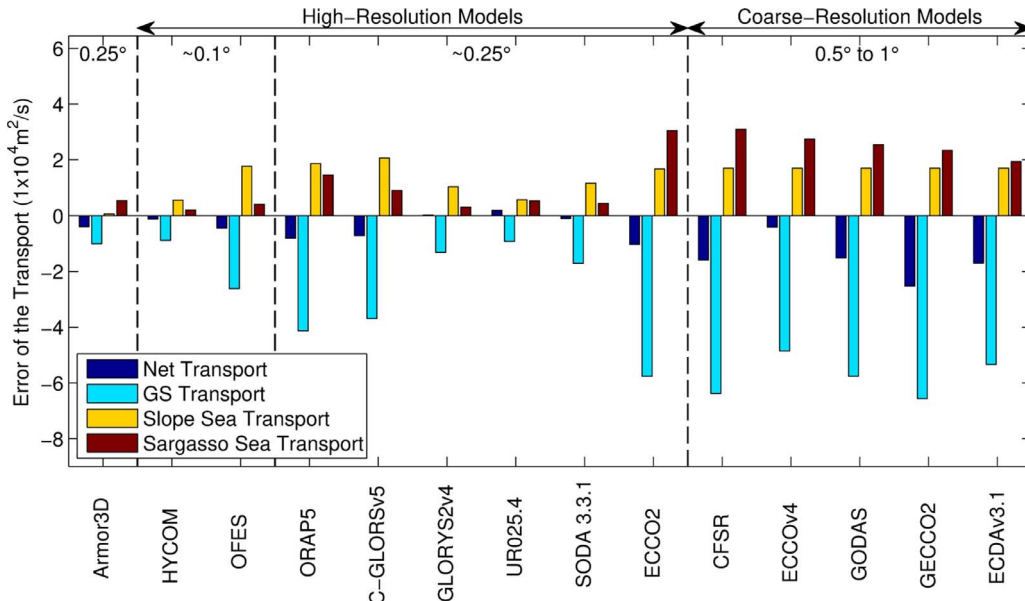


Fig. 10. Error of the net transport, Gulf Stream transport, Slope Sea transport and Sargasso Sea transport across the Oleander Line in different products.

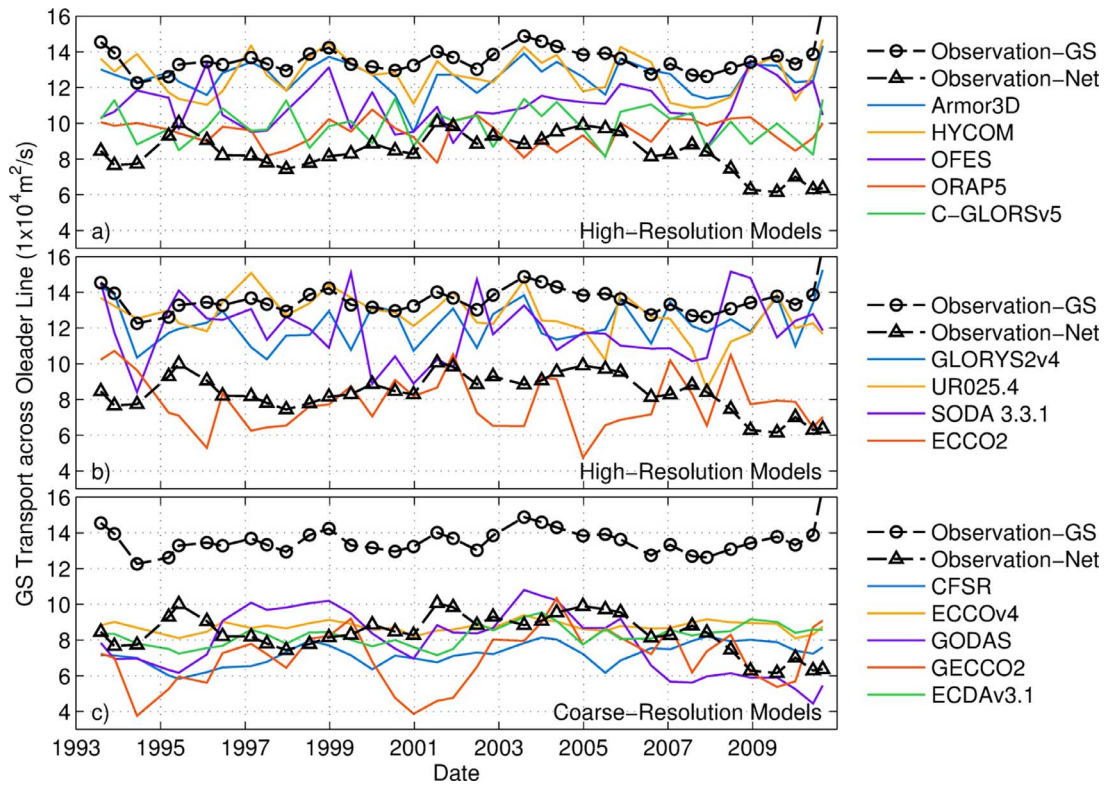


Fig. 11. Six-month running mean GS transport across the Oleander Line.

between 75°W and 50°W is close to the observed path (Fig. 14a and b). The GS path from ORAP5 overshoots northward after Cape Hatteras (between 74°W and 70°W), which indicates that ORAP5 does not reproduce the separation from the coast adequately. Considering ORAP5, C-GLORSv5 and UR025.4 are all based on NEMO, use the same

horizontal grid, similar forcing and assimilation methods, the overshooting problem in ORAP5 is probably not due to its grid resolution or assimilation method. Both ORAP5 and C-GLORSv5 have Slope Sea transports which are in the opposite direction as observed (Fig. 9), which could contribute to the GS overshoot (Zhang and Vallis, 2007);

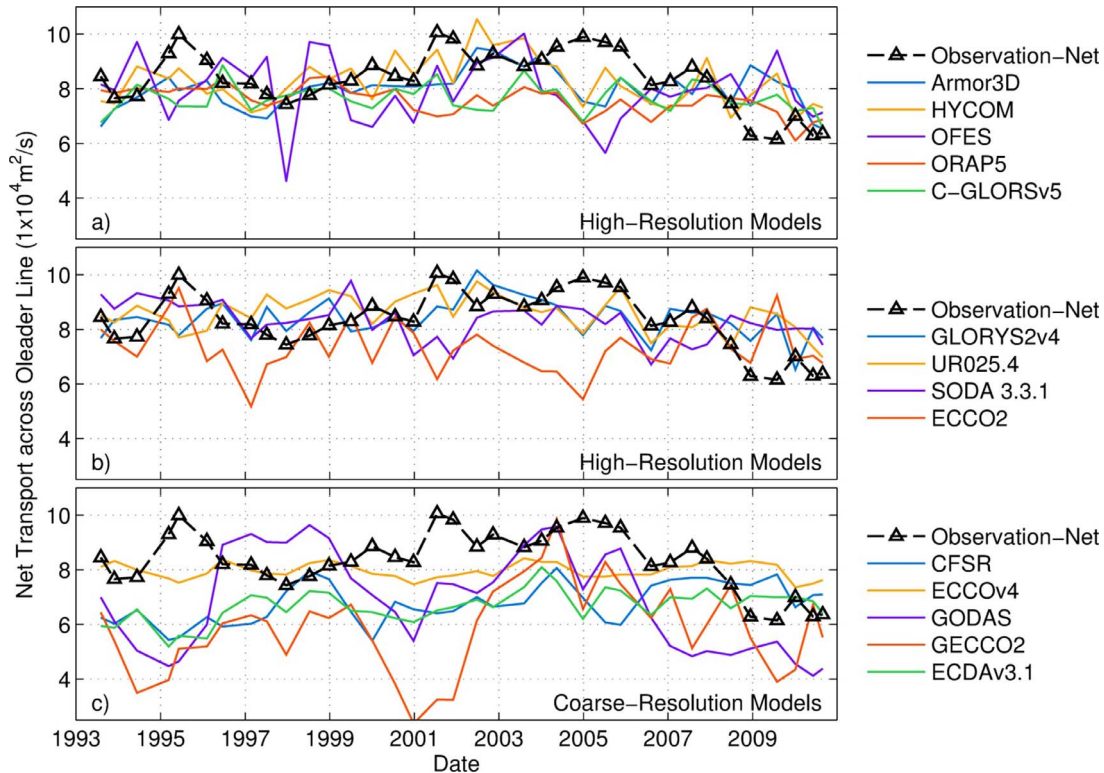


Fig. 12. Six-month running mean net transport across the Oleander Line.

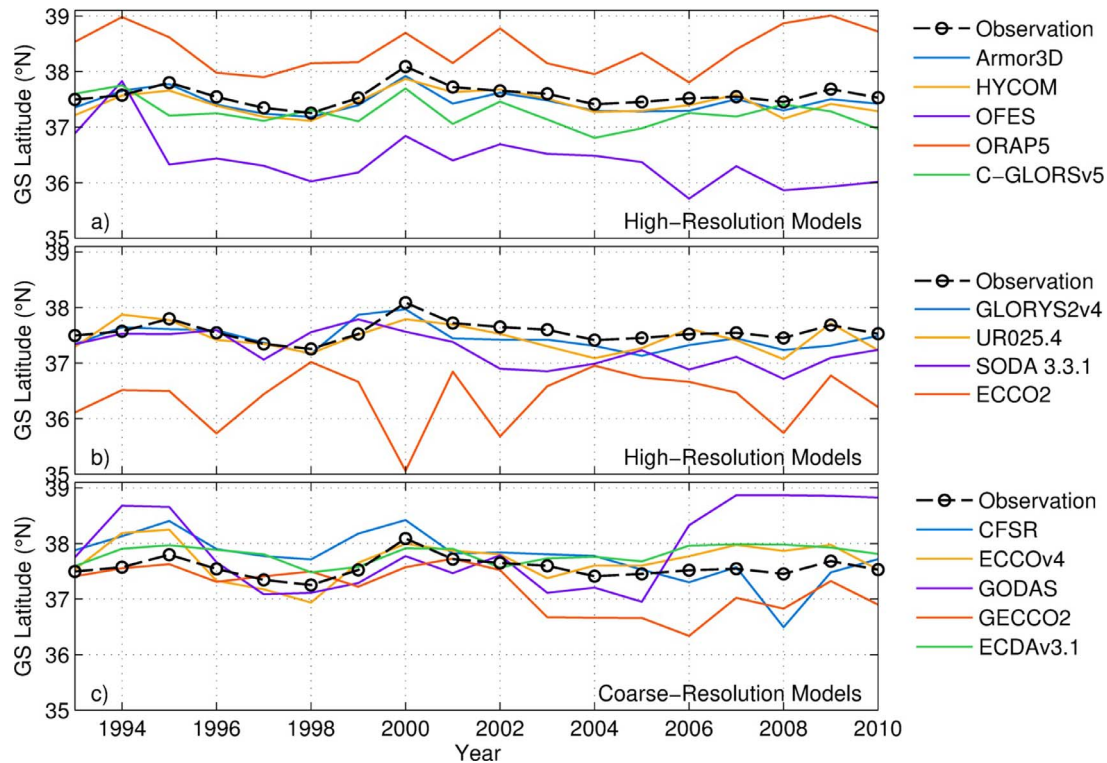


Fig. 13. Annual mean GS latitude at the Oleander Line.

however, the overshoot in C-GLORSv5 is much smaller than in ORAP5 even though its Slope Sea transport is more in error than ORAP5. A possible source of the overshooting GS path in ORAP5 is its sub-grid scale parameterization methods. The unrealistically large undulations throughout the GS path in ORAP5 indicates errors in the potential vorticity (PV) balance of the GS, which is sensitive to the details of the sub-grid scale parameterization.

In the two unconstrained models, OFES and ECCO2, the GS paths depart from the observed path at most longitudes. The GS paths in all coarse-resolution models (Fig. 14c) are biased between 75°W and 70°W, and show fewer undulations east of 70°W. The bias between 75°W and 70°W may due to the unrealistic topography in the coarse-resolution models.

In order to quantify model bias in the GS path, we calculated the absolute error area for the GS path (AEA-GS), which is the area between monthly snapshots of the GS path in numerical models and Aviso-ADT from 73°W to 55°W. An example of the AEA-GS calculation is shown in Fig. 15. The mean length of the GS path between 73°W and 55°W is about 2140 km, so a ~ 47 km meridional bias of the GS path results in $10 \times 10^{10} \text{ m}^2$ AEA-GS. Since the SSH in Armor3D is based on satellite observations, its mean AEA-GS value provides an estimate of errors due to interpolation.

HYCOM has the minimum mean AEA-GS among the numerical models (Fig. 16). On average, the GS path in HYCOM is biased by about 0.25°. GLORYS2v4 and UR025.4 show the second and third minimum mean AEA-GS. Not surprisingly, the two unconstrained models OFES and ECCO2 show the largest mean AEA-GS. The mean AEA-GS from all the other models are in a range from a little below $1 \times 10^{11} \text{ m}^2$ (ECCOv4) to $1.75 \times 10^{11} \text{ m}^2$ (GECCO2). The AEA-GSs in coarse-resolution models are comparable to that in some of the high-resolution models. Note that the AEA-GS is an estimate of the GS path bias between 73°W and 55°W and does not reflect the error in separation latitude; that is, a smaller AEA-GS does not necessarily imply a more accurate GS separation position at Cape Hatteras.

3.4.2. North Wall position

The northern edge of the GS is referred as the north wall for its strong horizontal temperature gradient persisting to great depth. It is occasionally visible in low-level atmospheric features due to strong local air-sea interaction (Sweet et al., 1981; Young and Sikora 2003). The meridional position of the north wall has proven to be important not only in local climate (Joyce et al., 2009; Kwon and Joyce 2013; Taylor 1996), but also in marine biology (Nye et al., 2011; Saba et al., 2015).

A number of operational definitions for the position of the north wall have been proposed, the most common being (1) the 15 °C isotherm at 200 m (Joyce et al., 2000; Joyce et al., 2009); (2) the 12 °C isotherm at 400 m (Bryan et al., 2007); (3) the location of SST front by IR satellite (Taylor and Stephens 1998); and (4) the location of temperature 2 °C lower than the maximum temperature in the current (Rossby and Gottlieb 1998). We use the 15 °C isotherm at 200 m to identify the north wall position in this paper, as the location of the 15 °C is readily obtained from hydrographic measurements going back several decades. We regard the 15 °C isotherm at 200 m from EN4 (dashed orange line in Fig. 1) to be the “observed” north wall position.

The biases of the mean north wall position (Fig. 17) from numerical models are similar to their biases of the mean GS path (Fig. 14). The topography changes dramatically near Cape Hatteras (Fig. 1) from more than 3000 m southeast of the north wall position to less than 100 m northwest of the north wall at (75°W, 36°N). Due to its coarse resolution, EN4 fails to recognize this sharp topography change. The north wall path from EN4 at 75°W, which is derived from 15 °C isotherm at 200 m, is located at a position shallower than 200 m. This is why the north wall in most high-resolution models is south to its position in EN4 at 75°W. The discrepancy of the north wall position close to its west end among coarse-resolution models is also partly due to their poor resolution of topography.

The overall bias of the north wall position in each model is estimated by the area between monthly snapshots of the GS north wall position derived from the model and EN4 from 73°W to 55°W, or absolute error area for north wall position (AEA-NW). The mean AEA-NW

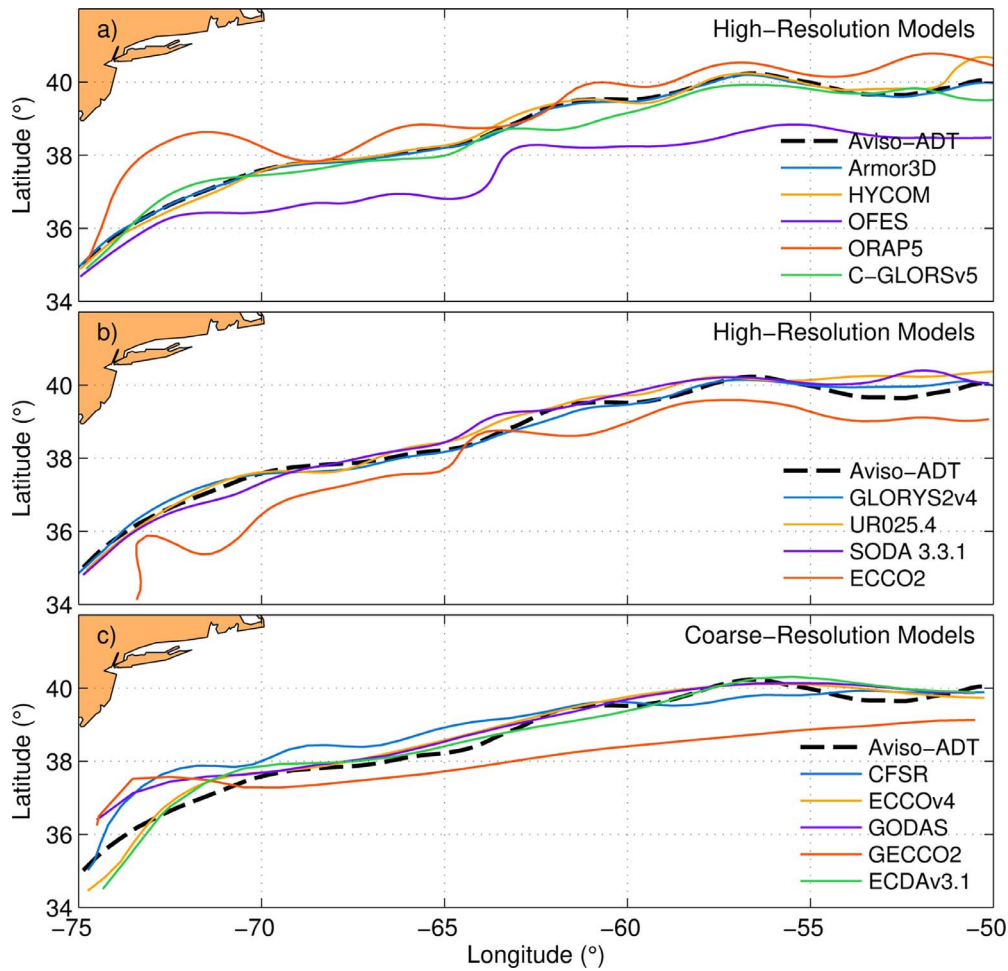


Fig. 14. 1993–2010 mean GS path from different datasets.

during 1993–2010 in each dataset is shown in Fig. 18. ECCOv4, with 1° horizontal resolution, has the smallest AEA-NW, which is smaller than the bias in the observationally based dataset Armor3D. The AEA-NW in HYCOM, C-GLORSv5, GLORYS2v4, UR025.4 and ECDav3.1 are slightly larger than that in Armor3D. Note that smaller error in ECCOv4 may partly be due to the horizontal resolution of the reference dataset (EN4), which is the same as in ECCOv4.

Several different indices have been introduced to describe the meridional position of the north wall as a whole. One which is well suited to model-data comparison is the Gulf Stream index (GSI) introduced by Joyce et al. (2000) and Joyce et al. (2009). It is defined as

the first principal component of 200 m temperature at nine points along the mean north wall position (shown in Fig. 1). The loading pattern of the first EOF is positive at all the nine points (Joyce et al., 2009), so positive (negative) values of its principal component indicate a coherent northward (southward) shift of the north wall.

The original GSI is derived from sparse in-situ temperature profiles from the World Ocean Database (WOD). Joyce et al. (2009) first interpolated 200 m temperature into $1^\circ \times 1^\circ$ grids in the middle of each season, where the temperature at each grid is the weighted average of all available temperature within $\pm 1^\circ$ in time, $\pm 1^\circ$ in latitude and $\pm 2^\circ$ in longitude. A Gaussian taper is applied during the

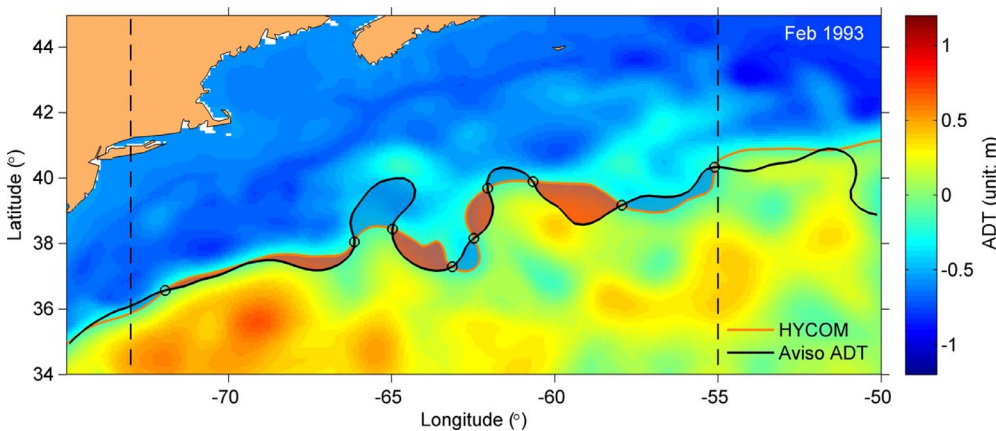


Fig. 15. An example of how the area between a reanalyzed SSH contour and the observed GS path (0.29-m ADT from satellite altimeters) is calculated. The solid black line is GS path from Aviso $\frac{1}{4}^\circ$ gridded ADT and the orange line is an SSH contour from HYCOM. The shading indicates regions between the two curves from 73°W to 55°W (black dashed lines). The red (blue) shading shows areas where the specific HYCOM SSH contour is north (south) to the observed GS path. The total area of both red and blue shaded regions is the absolute error area between the SSH contour from HYCOM and observed GS path. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

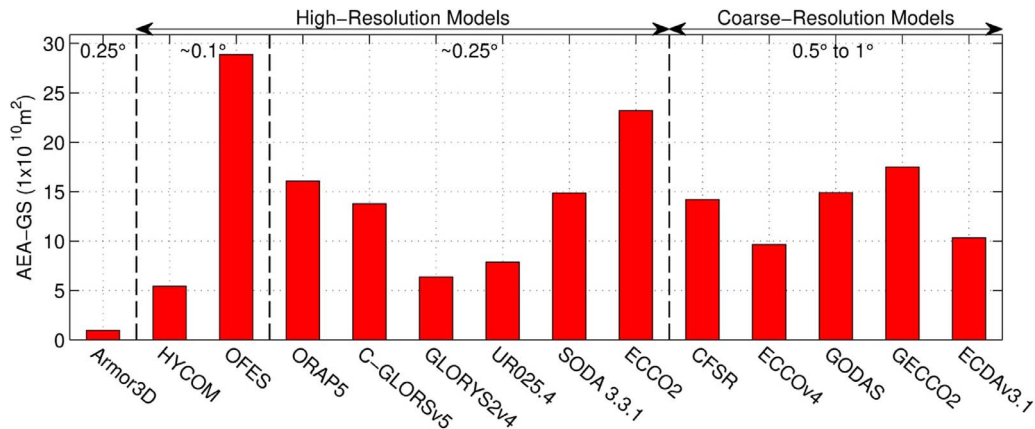


Fig. 16. Mean absolute error area (mean AEA-GS) of the GS path in each dataset during 1993–2010. $10 \times 10^{10} \text{ m}^2$ AEA-GS is equivalent to 46.7 km GS path bias in meridional direction. The horizontal resolution of the numerical models gets coarser from left to right and is marked on the top of the figure.

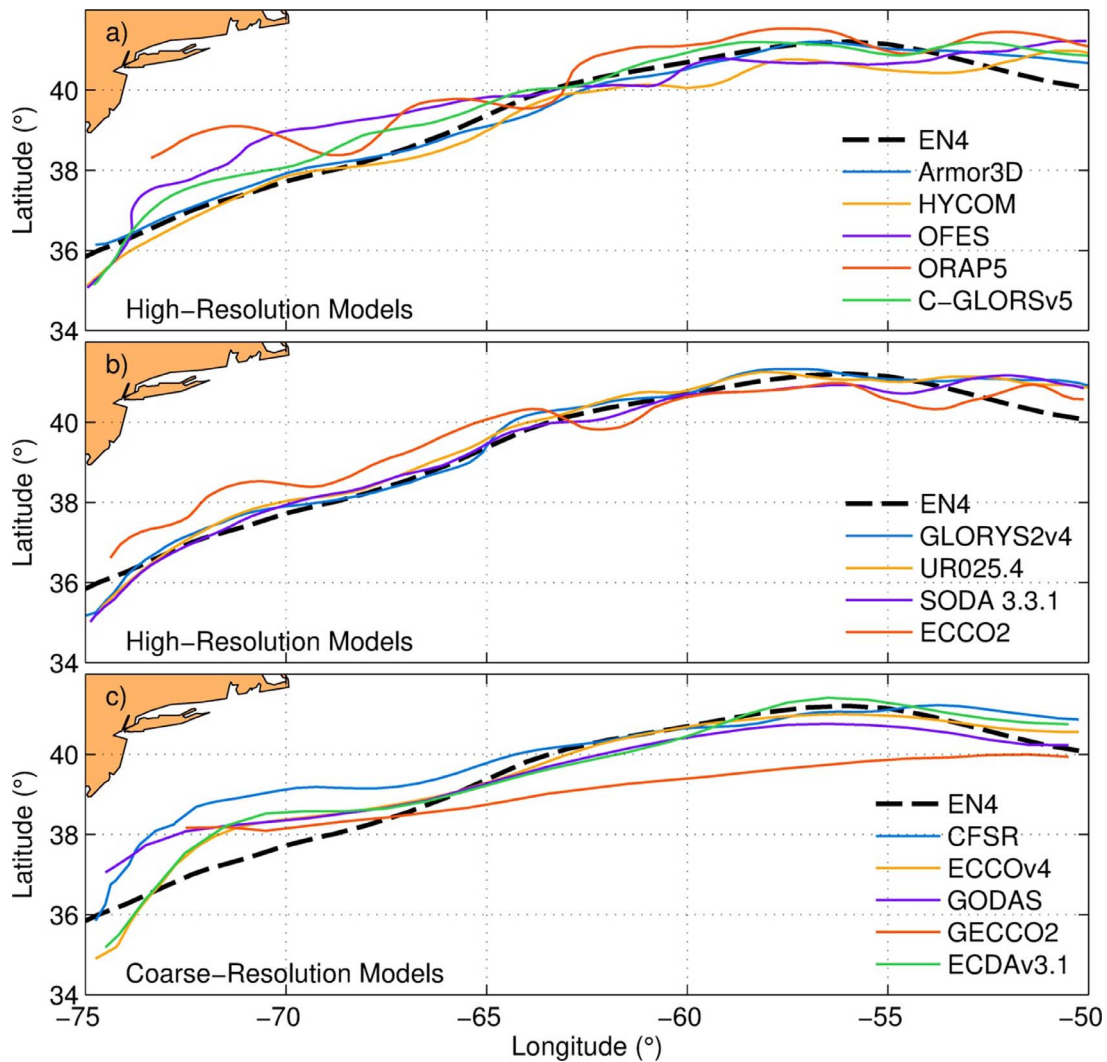


Fig. 17. The climate mean Gulf Stream northern wall positions from the 15 °C isotherms at 200 m during 1993–2010.

interpolation with e^2 -folding scale of 0.5 year in time, 1° in latitude and 2° in longitude. Finally, temperature at the nine selected points are interpolated from the gridded temperature.

For the comparison in this paper, we calculated a modified GSI (GSI-E8) with two major differences from the GSI introduced by Joyce et al. (2009). First, only the eastern 8 of the 9 locations are used since the

actual water depth at the west-most point used by Joyce et al. (2009) is less than 200 m. Second, the temperature at the eight points are interpolated from the seasonal mean temperature provided by EN4 (no Gaussian taper is applied). Detrended and normalized annual mean GSI and GSI-E8 are shown in Fig. 19. The annual mean GSI-E8 follows GSI well with a correlation coefficient of 79.6% at 99% confidence level.

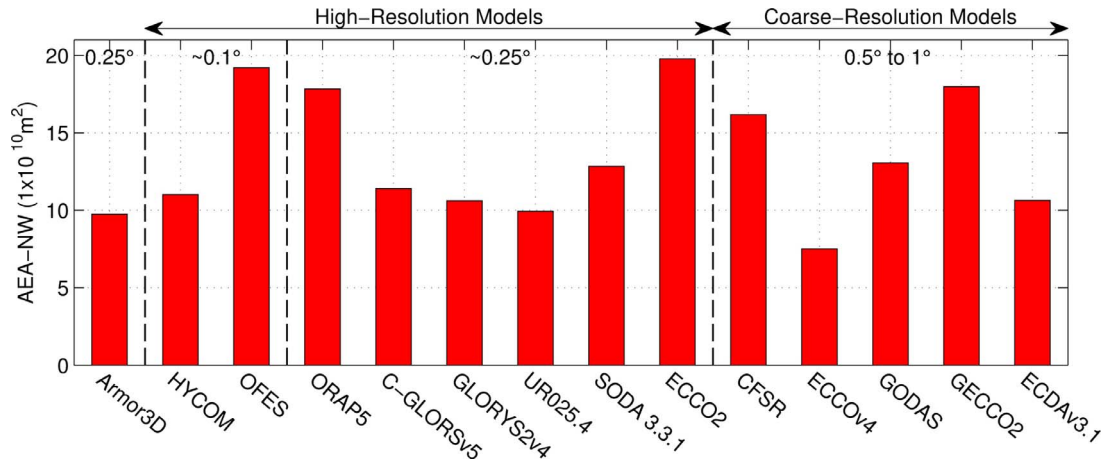


Fig. 18. Mean absolute error area of the GS north wall position (AEA-NW) in different datasets during 1993–2010. The horizontal resolution of the numerical models gets coarser from left to right and is marked on the top of the figure.

The PC of EOF depends sensitively on the analysis period of data, and only 18 years of data (1993–2010) can be obtained from the numerical models, which is much shorter than the interval used to calculate the GSI-E8 (1954–2012). Hence, to keep the reconstructed GSI-E8 (RGS-E8) derived from Armor3D and numerical models consistent with the GSI-E8 derived from EN4, RGS-E8 is derived via a fixed loading pattern, which is the loading pattern of first EOF mode during the calculation of GSI-E8 from EN4. Annual mean RGS-E8 (detrended and normalized) is shown in Fig. 20. The correlation coefficients between annual RGS-E8 and GSI-E8 are significant at 95% confidence level in 9 of the 13 numerical models: HYCOM (79%), C-GLORSv5 (83%), GLORYS2v4 (64%), UR025.4 (80%), SODA 3.3.1 (66%), CFSR (79%), GODAS (81%), ECCOv4 (65%), ECDAv3.1 (62%), and the observationally dataset, Armor3D (87%).

4. Discussion

4.1. The cross-stream GS structure

The skewness and kurtosis of the observed GS can be explained by noting that the GS is a PV as well as a temperature front, separating the high PV waters of the Slope Sea to the north from the low PV waters of the Sargasso Sea to the south. The PV increases rapidly as the core of the GS is crossed to the north, either at constant depth or while following a surface of constant density (Todd et al., 2016). A simple model of the GS can be formulated as an ideal PV front using the reduced gravity equations. We idealize the GS as a thin jet and, following Cushman-Roisin et al. (1993), make use of the reduced gravity equations on an f -plane. For the present case, it is only necessary to consider a single moving surface layer.

The PV in the moving layer is given by

$$q = \frac{-u_y + f}{h}, \quad (1)$$

where u is the downstream velocity, y is the cross-stream distance, h is the thickness of the moving layer, and subscripts indicate partial differentiation. We take $y = 0$ as the location of the PV front. For an ideal front, the PV distribution is

$$q(y) = \begin{cases} q_s & y < 0, \\ q_n & y > 0, \end{cases} \quad (2)$$

with $q_n > q_s$. Since the zonal velocity vanishes far from the front, the two values of the PV are

$$q_s = \frac{f}{H_s} \text{ and } q_n = \frac{f}{H_n},$$

where H_s and H_n are interpreted as the depths of the thermocline far to the south and north, respectively, of the GS.⁸ Eqs. (1) and (2) provide a relationship between the layer thickness, velocity, and the conditions far from the front. The system is closed by the geostrophic relation

$$u = -\frac{g'}{f} h_y, \quad (3)$$

where g' is the reduced gravity. Inserting (3) into (1) and (2) gives a differential equation in h :

$$h_{yy} - \frac{fq(y)}{g'} h = -\frac{f^2}{g'}. \quad (4)$$

The solution to (4) for the ideal PV distribution (2) gives this downstream velocity profile

$$u = f(L_s - L_n) \begin{cases} e^{\frac{y}{L_s}} & y < 0, \\ e^{-\frac{y}{L_n}} & y > 0, \end{cases} \quad (5)$$

where

$$L_{s,n} = \frac{\sqrt{g'H_{s,n}}}{f}$$

are the Rossby deformation radii on the south and north side of the GS. An important feature of the solution (4) is its asymmetry about the origin—the velocity decays away from the core of the GS more slowly to the south than to the north because the Rossby radius is larger on the southern side of the GS than on the northern side. The velocity profile (4) is compared to the Oleander velocity observations at 55 m in Fig. 21. The only adjustable parameters in (4) are the two Rossby radii; these were determined by a least-square fit to the Oleander observations over the range where the observed velocity is positive. The optimal radii, $L_s = 46$ km and $L_n = 22$ km, are broadly consistent with the observed Rossby radii in the Sargasso and Slope Seas (e.g., Chelton et al., 1998). The RMS error in the fit is 8%—remarkably small for such a simple model.

The observed velocity profile is smoother than the ideal-front velocity profile and the velocity profiles obtained from the numerical models are smoother still. Any amount of turbulence or numerical dissipation will erode a sharp front like that given in (2) resulting in a smoother PV distribution and, as a result, a smoother velocity profile. A crude model of diffusive spreading of the PV front can be constructed by imagining that the ideal front (2) is representative of the GS

⁸ The downstream direction is inclined with respect to due east at the Oleander Line, so “north” and “south” in this context are more precisely “northwest” and “southeast”.

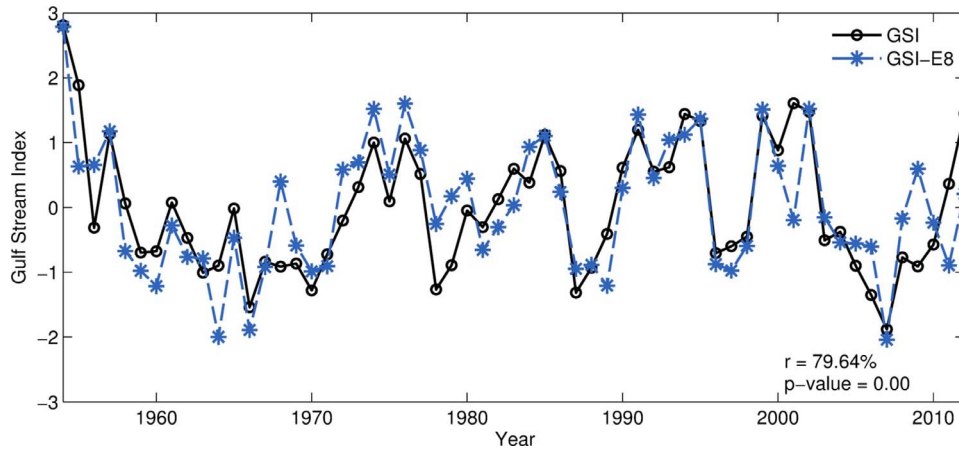


Fig. 19. Annual mean Gulf Stream Index (GSI) from Joyce et al. (2009) (solid line with circles) and reproduced GSI (GSI-E8) from EN4 based on the eastern 8 locations (dashed line with asterisks). Both indexes are normalized and detrended.

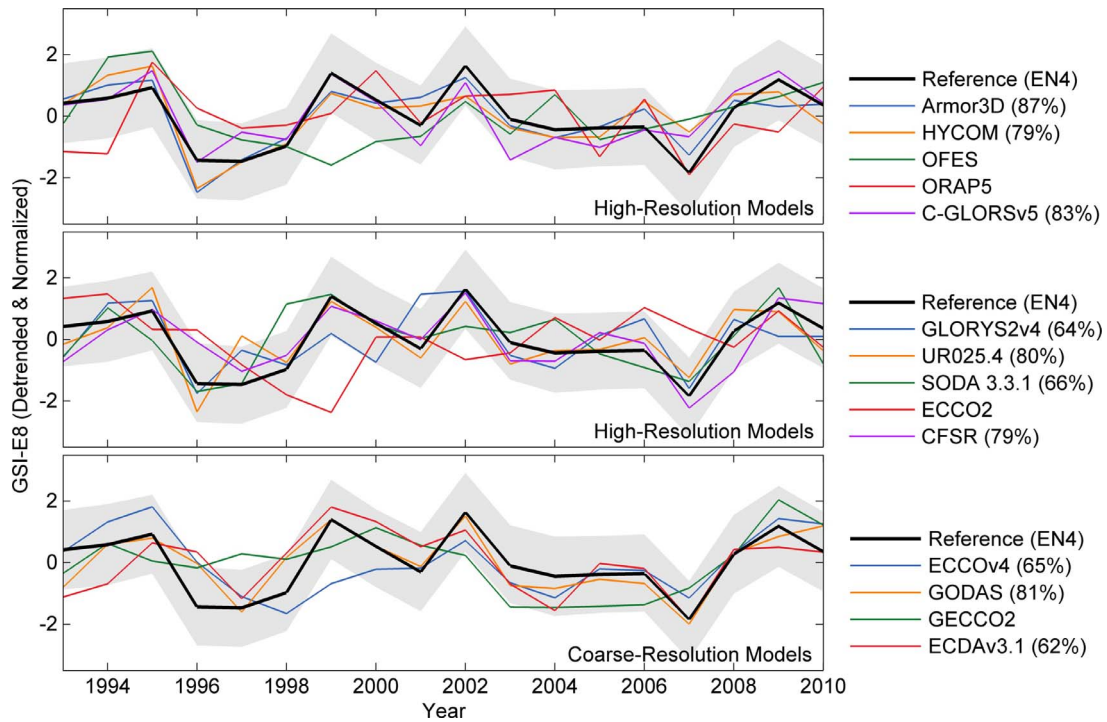


Fig. 20. the GSI-E8 derived from EN4 and reconstructed GSI (RGS-E8) from Armor3D and reanalysis products. The shaded region is one standard deviation of GSI-E8 uncertainty from EN4. The numbers in the legend are the correlation coefficients between the RGS-E8 from reanalyses and GSI-E8 from EN4. The insignificant correlation coefficients (p -value > 0.05) are not shown.

structure immediately after separation at Cape Hatteras; thereafter, the PV front is advected downstream and diffused laterally. In steady state, the PV budget is

$$\bar{u}q_x = K_e q_{yy} \quad (6)$$

where x is the downstream coordinate, K_e is an eddy diffusivity of PV, and $\bar{u}$ is the mean downstream velocity which, for simplicity, we consider to be independent of the cross- and downstream coordinates. The solution to (6) with upstream conditions (2) is

$$q = \frac{f}{2} \left(\frac{1}{H_n} + \frac{1}{H_s} \right) + \frac{f}{2} \left(\frac{1}{H_n} - \frac{1}{H_s} \right) \operatorname{erf} \left(\frac{y}{\sqrt{2}d} \right), \quad (7)$$

where

$$d = \sqrt{\frac{2K_e x}{\bar{u}}}$$

and x is the downstream distance from Cape Hatteras to the Oleander Line (about 600 km). Note that the PV distribution (7) is the same as that obtained by convolving the ideal-front distribution (2) with a Gaussian with standard deviation d .

The equation for the moving layer thickness (4) is challenging to solve analytically with the PV distribution (7), but it is readily solved numerically by finite differences. The solutions presented here were obtained by discretizing (7) on a domain of width $10L_s$ with a grid spacing of $0.001L_s$. As with the ideal-front model, the Rossby radii and smoothing width are found by a least-squares fit to the Oleander observations. The best fit, shown in Fig. 21, has Rossby radii $L_s = 43$ km and $L_n = 17$ km, and a smoothing width of $d = 10$ km. Note that the maximum velocity obtained from the smoothed-front model is displaced toward the low-PV side (i.e., the south side); the profile in Fig. 21 has been shifted horizontally so the maximum velocity is again found at $y = 0$. The optimal Rossby radii are not significantly different

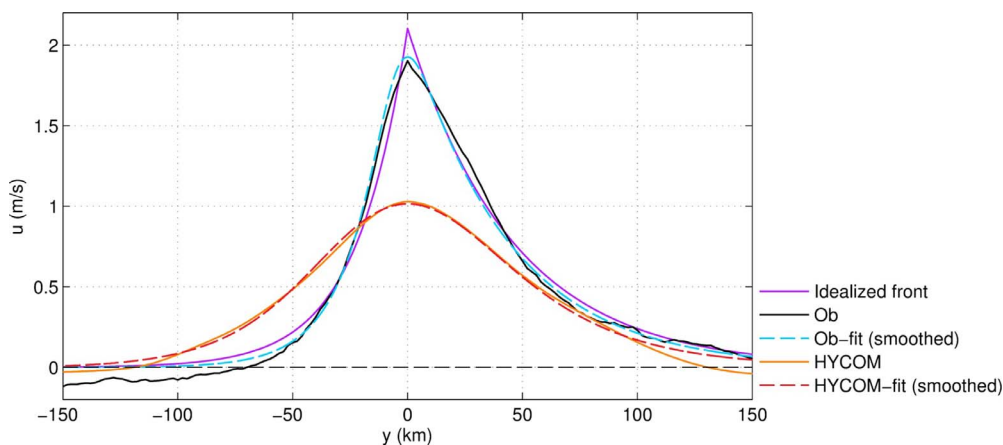


Fig. 21. GS at 55 m velocity in m s^{-1} from the Oleander Project (black solid line), ideal front (pink solid line), smoothed-front (cyan dashed line) models as a function of cross stream distance in km. The axis is oriented as if the viewer were looking downstream, so y increases toward the south. Also shown are the velocity profile obtained from the HYCOM product (orange solid line) and a smoothed-front fit to the HYCOM profile (red dashed line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

from the ideal-front model. If we use 600 km and 1 m s^{-1} for the distance to Cape Hatteras and the advecting velocity, respectively, $d = 10 \text{ km}$ corresponds to a cross-stream eddy diffusivity for PV of $K_e \approx 166 \text{ m}^2 \text{ s}^{-1}$. The smoothed-front model fits the observations slightly better than the ideal-front model, with an RMS error of 5%.

The smoothed-front model is fit to the remaining observational and high-resolution products; the PV front is too diffuse in the coarse-resolution models to reliably estimate the required fitting parameters. The fit to the HYCOM velocity profile is shown in Fig. 21. The best-fit Rossby radii are similar to those obtained from the Oleander observations ($L_s = 38 \text{ km}$ and $L_n = 11 \text{ km}$), but the smoothing width is significantly larger at $d = 58 \text{ km}$. The equivalent cross-front diffusivity is $K_e \approx 5500 \text{ m}^2 \text{ s}^{-1}$, about thirty-three times larger than that inferred from observations. The PV front in the HYCOM product is therefore significantly broader and smoother than that found from observations. This indicates that the subgrid-scale parameterizations of dissipation used by HYCOM are unrealistically efficient at mixing PV across the GS, the core of which has been found to be a barrier to mixing by numerous studies (e.g., Bower et al., 1985; Bower and Rossby 1989; Bower 1991; Samelson 1992; Bower and Lozier 1994; Lozier et al., 1997). The implication is that, even at $1/12^\circ$ resolution, HYCOM's dynamics are simply too diffuse to properly represent the sharpness of the GS PV front. Data assimilation cannot force the model to maintain a PV front which is sharper than its dynamics allow, although it may prevent the front from spreading beyond this fundamental limit.

The results obtained from fitting the smoothed-front model to the other high-resolution products (and Armor3D) are summarized in Fig. 22. The misfit is always less than 11% and the Rossby radii are all within a physically reasonable range, though with some variation. With the exception of the Oleander observations, the smoothing widths are all about 1–1.5 times the southern Rossby radius and the implied PV diffusivities are nearly two orders of magnitude larger than that inferred from the observed velocity profile. Armor3D is a gridded, but observationally-based, data set and therefore does not have the PV mixing dynamics of the numerical models. In this case, the smoothing kernels inherent in the optimal interpolation procedure play a role similar to PV mixing; the inferred PV diffusivity could thus be thought of as an effective diffusivity induced by the interpolation.

When the smoothing width exceeds the Rossby radii on the northern and southern flanks of the GS, the scale of the velocity's decay from its maximum value is dominated by the smoothing scale. The velocity profiles obtained from the gridded dataset and numerical products are thus broader and more symmetric than the observed velocity profile. For a given transport, the peak GS velocity must decrease as the GS becomes broader. As all of the numerical products have excessively broad GSs, they all underestimate the maximum GS velocity. Products which also significantly underestimate the GS transport have even smaller peak velocities.

4.2. Overall performance of the numerical models

The performance of each dataset discussed in Section 3 is summarized in Table 2. To show the overall performance of each dataset more clearly, the models closest to observations in each aspect is marked in bold and the four models closest to the observations are marked by shaded backgrounds in the table. Two of the models, HYCOM and GLORYS2v4, are among the best for nearly all aspects. HYCOM is among the top 4 models in 21 of the 25 listed aspects in Table 2. And GLORYS2v4 is among the top 4 numerical models in 20 of the 25 aspects. Generally, high-resolution models perform better than coarse-resolution models. The noteworthy performance of HYCOM might be expected since it is the highest-resolution constrained product discussed in this paper; however, it is not *a priori* clear that assimilation of sparse observational data should lead to similar improvements an eddy-resolving model as seen in more coarsely resolved models. In the case of HYCOM, it appears that data assimilation continues to yield dividends at high resolution. In contrast, the superiority of GLORYS2v4 over ORAP5, C-GLORYSv5 and UR025.4 is surprising because all the four constrained models are based on NEMO, sharing same horizontal grid and similar forcing. One possible reason is that GLORYS2v4 is constrained via a reduced order Kalman filter method,⁹ while ORAP5, C-GLORYSv5 and UR025.4 are constrained via 3DVAR method. The Kalman filter method has been shown to give better performance than 3DVAR in many cases (Miyoshi, 2005; Whitaker et al., 2008; Yang et al., 2009).

5. Summary

An observationally based 3D ocean dataset (Armor3D) and thirteen numerical models, including eleven ocean reanalysis products (HYCOM, ORAP5, C-GLORYSv5, GLORYS2v4, UR025.4, SODA 3.3.1, CFSR, ECCOV4, GODAS, GECCO2, ECDav3.1) and two unconstrained ocean simulations (OFES, ECCO2), are compared with in-situ, analyzed, and remote observations. The intercomparison is focused on the Gulf Stream from the Florida Straits to south of the Grand Banks and considers the Florida Current transport, the separation of Gulf Stream at Cape Hatteras, Gulf Stream transport, position and structure along the Oleander Line, Gulf Stream path, and the north wall of the Gulf Stream. The variations between products in the cross-stream structure of the GS is explained using a simple model of the GS as a smoothed PV front. Broader PV fronts lead to broader and more symmetric velocity profiles as well as lower peak velocities. It is suggested that the failure of the numerical models to reproduce the observed peak GS velocity is due to

⁹ The 3DVAR method is also applied to correct the large-scale temperature and salinity distribution in GLORYS2v4 when enough observations are available (Garric and Parent, 2017).

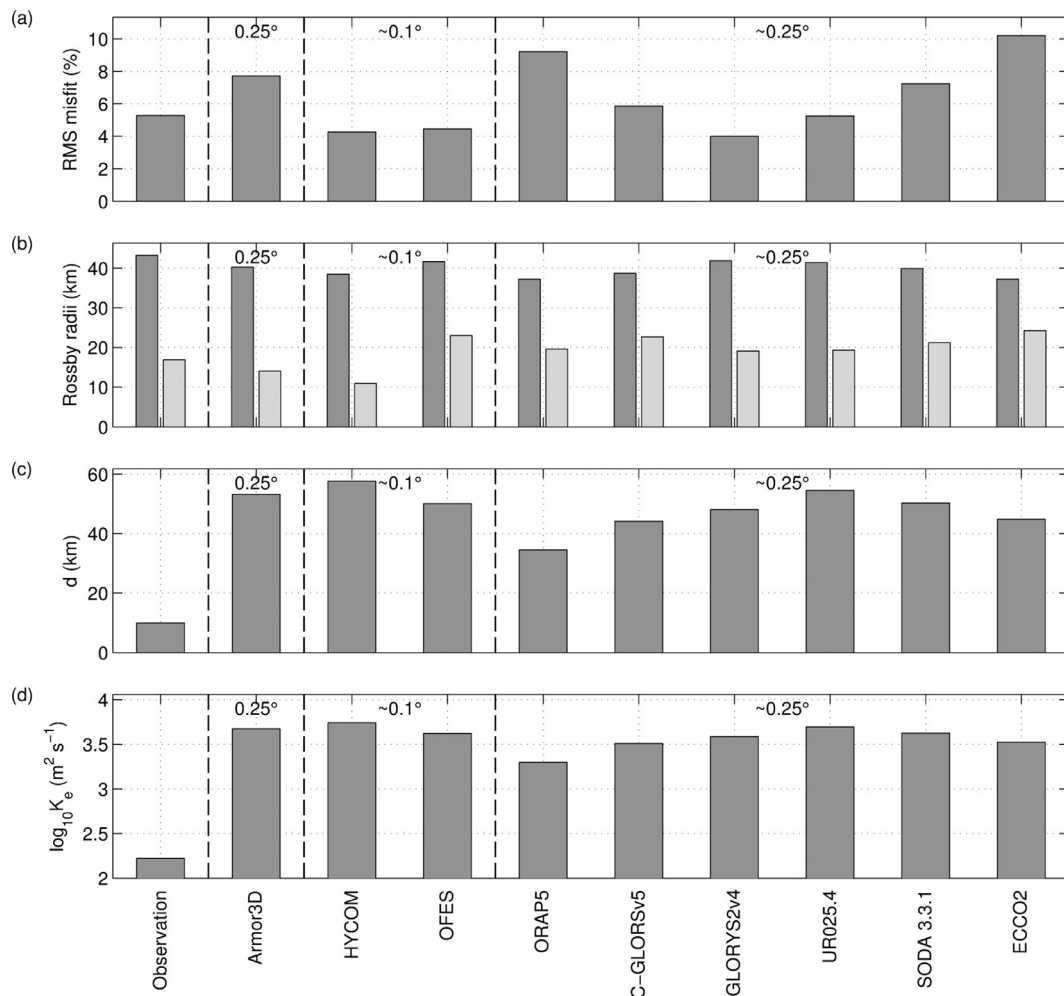


Fig. 22. Summary statistics for the smoothed-front model fit to observations and numerical products. The statistics are: (a) the RMS misfit between the smoothed-front model and the observed or model-generated velocity profile; (b) the best-fit Rossby radii in km, with the southern and northern radii in dark and light gray, respectively; (c) the best-fit smoothing distance in km; and (d) the eddy diffusivity for PV in $m^2 s^{-1}$ implied by the smoothing distance given in (c).

overly-aggressive sub-grid scale dissipation which erodes the GS PV front.

Among the 13 numerical models, the GS in two of them, HYCOM and GLORYS2v4, is closest to the observations. HYCOM's primary advantage is likely its high resolution (1/120); the good performance of GLORYS2v4 may be due to its assimilation method (Kalman filter).

Even though reanalysis products are expected to improve numerical simulations and provide near-observation results, they do not have to be so. We show in this paper that, in fact, most available reanalysis products fail to adequately represent some of the most basic features of the Gulf Stream. Excluding the two free-run simulations, OFES and ECCO2, less than half of the 11 constrained products are significantly (95% level) correlated with the observations in the Gulf Stream separation latitude, the Gulf Stream and the net transport along the Oleander Line. In some of the models, the 11-year mean values may also be largely biased, such as the Gulf Stream transport across the Oleander Line. In addition, some common errors exist in the numerical models discussed in the paper. None of the models can reproduce the correct horizontal structure and maximum value of the velocity along the Oleander Line. There is thus much work to be done before ocean reanalyses can be regarded as approximations to reality with the same confidence commonly attached to atmospheric reanalyses.

Acknowledgment

This work was supported by the National Science Foundation (OCE-

1634829). The authors thank Terrence M. Joyce for providing GSI data and two anonymous reviewers for providing comments which improved the manuscript.

References

- Allard, R., Metzger, E.J., Broome, R., Franklin, D., Smedstad, O.M., Wallcraft, A., 2013. Development of an eddy-resolving reanalysis using the 1/12° global HYbrid coordinate ocean model and the navy coupled ocean data assimilation scheme. In: *Proceedings of the EGU General Assembly Conference Abstracts*. Vienna, Austria. pp. 7–12 Apr.
- Andres, M., 2016. On the recent destabilization of the gulf stream path downstream of cape hatteras. *Geophys. Res. Lett.* 43, 9836–9842. <http://dx.doi.org/10.1002/2016GL069966>.
- Armour, K.C., Marshall, J., Scott, J.R., Donohoe, A., Newsom, E.R., 2016. Southern ocean warming delayed by circumpolar upwelling and equatorward transport. *Nat. Geosci.* 9, 549–554. <http://dx.doi.org/10.1038/ngeo2731>.
- Balmaseda, M.A., Hernandez, F., Storto, A., Palmer, M.D., Alves, O., Shi, L., Smith, G.C., Toyoda, T., Valdivieso, M., Barnier, B., Behringer, D., Boyer, T., Chang, Y.-S., Chepurin, G.A., Ferry, N., Forget, G., Fujii, Y., Good, S., Guinehut, S., Haines, K., Ishikawa, Y., Keeley, S., Köhl, A., Lee, T., Martin, M.J., Masina, S., Masuda, S., Meyssignac, B., Mogensén, K., Parent, L., Peterson, K.A., Tang, Y.M., Yin, Y., Vernieres, G., Wang, X., Waters, J., Wedd, R., Wang, O., Xue, Y., Chevallier, M., Lemieux, J.-F., Dupont, F., Kuragano, T., Kamachi, M., Awaji, T., Caltabiano, A., Wilmer-Becker, K., Gaillard, F., 2015. The ocean reanalyses intercomparison project (ORA-IP). *J. Oper. Oceanogr.* 8, s80–s97. <http://dx.doi.org/10.1080/1755876X.2015.1022329>.
- Behringer, D., Xue, Y., 2004. Evaluation of the global ocean data assimilation system at NCEP: the pacific ocean. In: *Proceedings of the Eighth Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, AMS 84th Annual Meeting. Seattle, Washington. Washington State Convention and Trade Center, pp. 11–15.

- Bower, A.S., 1991. A simple kinematic mechanism for mixing fluid parcels across a meandering jet. *J. Phys. Oceanogr.* 21, 173–180. [http://dx.doi.org/10.1175/1520-0485\(1991\)021<0173:ASKMFM>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1991)021<0173:ASKMFM>2.0.CO;2).
- Bower, A.S., Rossby, H.T., Lillibridge, J.L., 1985. The gulf stream—barrier or blender? *J. Phys. Oceanogr.* 15, 24–32. [http://dx.doi.org/10.1175/1520-0485\(1985\)015<0024:TGSOB>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1985)015<0024:TGSOB>2.0.CO;2).
- Bower, A.S., Rossby, T., 1989. Evidence of cross-frontal exchange processes in the gulf stream based on isopycnal RAFOS float data. *J. Phys. Oceanogr.* 19, 1177–1190. [http://dx.doi.org/10.1175/1520-0485\(1989\)019<1177:EOCFEP>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1989)019<1177:EOCFEP>2.0.CO;2).
- Bower, A.S., Lozier, M.S., 1994. A closer look at particle exchange in the gulf stream. *J. Phys. Oceanogr.* 24, 1399–1418. [http://dx.doi.org/10.1175/1520-0485\(1994\)024<1399:ACLAP>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1994)024<1399:ACLAP>2.0.CO;2).
- Bleck, R., 2002. An oceanic general circulation model framed in hybrid isopycnal-Cartesian coordinates. *Ocean Model* 4, 55–88. [http://dx.doi.org/10.1016/S1463-5003\(01\)00012-9](http://dx.doi.org/10.1016/S1463-5003(01)00012-9).
- Bryan, F.O., Hecht, M.W., Smith, R.D., 2007. Resolution convergence and sensitivity studies with North Atlantic circulation models. Part I: the western boundary current system. *Ocean Model* 16, 141–159. <http://dx.doi.org/10.1016/j.ocemod.2006.08.005>.
- Bryden, H.L., King, B.A., McCarthy, G.D., McDonagh, E.L., 2014. Impact of a 30% reduction in Atlantic meridional overturning during 2009–2010. *Ocean Sci.*, 10, 683–691. <http://dx.doi.org/10.5194/os-10-683-2014>.
- Chang, Y.S., Zhang, S., Rosati, A., Delworth, T.L., Stern, W.F., 2013. An assessment of oceanic variability for 1960–2010 from the GFDL ensemble coupled data assimilation. *Clim. Dyn.* 40, 775–803. <http://dx.doi.org/10.1007/s00382-012-1412-2>.
- Chassignet, E.P., Marshall, D.P., 2008. Gulf Stream separation in numerical ocean models. In: Hecht, M.W., Hasumi, H. (Eds.), *Ocean Modeling in an Eddying Regime*. In: Hecht, M.W., Hasumi, H. (Eds.), *Geophysical monograph series 177*. American Geophysical Union, pp. 39–61. <http://dx.doi.org/10.1029/177GM05>.
- Chassignet, E.P., Xu, X., 2017. Impact of horizontal resolution (1/12° to 1/50°) on Gulf stream separation, penetration, and variability. *J. Phys. Oceanogr.* 47, 1999–2021. <https://doi.org/10.1175/JPO-D-17-0031.1>.
- Chelton, D.B., deSzoeke, R.A., Schlax, M.G., Naggar, E.L., Siwertz, 1998. Geographical variability of the first baroclinic Rossby radius of deformation. *J. Phys. Oceanogr.* 28, 433–460. [http://dx.doi.org/10.1175/1520-0485\(1998\)028<0433:GVOTFB>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1998)028<0433:GVOTFB>2.0.CO;2).
- Carton, J.A., Chepurin, G.A., Chen, L., 2018. An updated reanalysis of ocean climate using the Simple Ocean Data Assimilation version 3 (SODA3), (manuscript in preparation), http://www.atmos.umd.edu/~ocean/index_files/soda3_readme.htm.
- Cornillon, P., 1986. The effect of the new England seamounts on gulf stream meandering as observed from satellite IR imagery. *J. Phys. Oceanogr.* 16, 386–389. [http://dx.doi.org/10.1175/1520-0485\(1986\)016<0386:TEOTNE>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1986)016<0386:TEOTNE>2.0.CO;2).
- Cushman-Roisin, B., Pratt, L., Ralph, E., 1993. A general theory for equivalent barotropic thin jets. *J. Phys. Oceanogr.* 23 (1), 91–103. [http://dx.doi.org/10.1175/1520-0485\(1993\)023<0091:AGTFEB>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1993)023<0091:AGTFEB>2.0.CO;2).
- Dieng, H.B., Palanisamy, H., Cazenave, A., Meyssignac, B., von Schuckmann, K., 2015. The sea level budget since 2003: inference on the deep ocean heat content. *Surv. Geophys.* 36, 209–229. <http://dx.doi.org/10.1007/s10712-015-9314-6>.
- DiNezio, P.N., Gramer, L.J., Johns, W.E., Meinen, C.S., Baringer, M.O., 2009. Observed interannual variability of the Florida current: wind forcing and the North Atlantic oscillation. *J. Phys. Oceanogr.* 39, 721–736. <http://dx.doi.org/10.1175/2008JPO4001.1>.
- Flagg, C.N., Schwartze, G., Gottlieb, E., Rossby, T., 1998. Operating an acoustic Doppler current profiler aboard a container vessel. *J. Atmos. Ocean. Technol.* 15, 257–271. [http://dx.doi.org/10.1175/1520-0426\(1998\)015<0257:OAADCP>2.0.CO;2](http://dx.doi.org/10.1175/1520-0426(1998)015<0257:OAADCP>2.0.CO;2).
- Forget, G., Campin, J.-M., Heimbach, P., Hill, C.N., Ponte, R.M., Wunsch, C.I., 2015. ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model Dev.* 8, 3071–3104. <http://dx.doi.org/10.5194/gmd-8-3071-2015>.
- Fuglister, F.C., Voorhis, A.D., 1965. A new method of tracking the Gulf Stream. *Limnol. Oceanogr.* 10, R115–R124. <http://dx.doi.org/10.4319/lo.1965.10.suppl2.r115>.
- Gangopadhyay, A., Cornillon, P.C., Watts, D.R., 1992. A test of the Parsons-Veronis hypothesis on the separation of the Gulf Stream. *J. Phys. Oceanogr.* 22, 1386–1399. [http://dx.doi.org/10.1175/1520-0485\(1992\)022<1386:ATOTPHS>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1992)022<1386:ATOTPHS>2.0.CO;2).
- Gangopadhyay, A., Chaudhuri, A.H., Taylor, A.H., 2016. On the nature of temporal variability of the Gulf stream path from 75° to 55°W. *Earth Interact.* 20, 1–17. <http://dx.doi.org/10.1175/EI-D-15-0025.1>.
- Garric, G., Parent, L., 2017. Product user manual for global ocean reanalysis products GLOBAL-REANALYSIS-PHY-001-025. <http://cmems-resources.cls.fr/documents/PUM/CMEMS-GLO-PUM-001-025-011-017.pdf> (accessed 8.4.17).
- Goni, G., Baringer, M., Bringas, F., 2017. Florida Current Transport in North Atlantic Ocean Derived from Satellite Altimetry Observations (NCEI Accession 0157747). NOAA National Centers for Environmental Information <http://dx.doi.org/10.7289/V55B00HT.v1.1>.
- Good, S.A., Martin, M.J., Rayner, N.A., 2013. EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res. Ocean.* 118, 6704–6716. <http://dx.doi.org/10.1002/2013JC009067>.
- Griffies, S.M., Harrison, M.J., Pacanowski, R.C., Rosati, A., 2004. A Technical Guide to MOM4. NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, USA, pp. 371.
- Guinehut, S., Dhomp, A.L., Larnicol, G., Le Traon, P.Y., 2012. High resolution 3-D temperature and salinity fields derived from in situ and satellite observations. *Ocean Sci.* 8, 845–857. <http://dx.doi.org/10.5194/os-8-845-2012>.
- Häkkinen, S., Rhines, P.B., Worthen, D.L., 2015. Heat content variability in the North Atlantic ocean in ocean reanalyses. *Geophys. Res. Lett.* 42, 2901–2909. <http://dx.doi.org/10.1002/2015GL063299>.
- Haines, K., Stepanov, V.N., Valdivieso, M., Zuo, H., 2013. Atlantic meridional heat transports in two ocean reanalyses evaluated against the RAPID array. *Geophys. Res. Lett.* 40, 343–348. <http://dx.doi.org/10.1029/2012GL054581>.
- Hurlburt, H.E., Hogan, P.J., 2000. Impact of 1/8° to 1/64° resolution on Gulf Stream model-data comparisons in basin-scale subtropical Atlantic Ocean models. *Dyn. Atmos. Ocean.* 32, 283–329. [http://dx.doi.org/10.1016/S0377-0265\(00\)00050-6](http://dx.doi.org/10.1016/S0377-0265(00)00050-6).
- Hurlburt, H.E., Metzger, E.J., Richman, J.G., Chassignet, E.P., Drillet, Y., Hecht, M.W., Le Galloudec, O., Shriver, J.F., Xu, X., Zambudio, L., 2011. Dynamical evaluation of ocean models using the Gulf Stream as an example. In: Schiller, A., Brassington, G.B. (Eds.), *Operational Oceanography in the 21st Century*. Springer, pp. 545–609.
- Joyce, T.M., Deser, C., Spall, M.A., 2000. The relation between decadal variability of subtropical mode water and the North Atlantic oscillation. *J. Clim.* 13, 2550–2569. [http://dx.doi.org/10.1175/1520-0442\(2000\)013<2550:TRBDVO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2000)013<2550:TRBDVO>2.0.CO;2).
- Joyce, T.M., Kwon, Y.O., Yu, L., 2009. On the relationship between synoptic wintertime atmospheric variability and path shifts in the Gulf stream and the Kuroshio extension. *J. Clim.* 22, 3177–3192. <http://dx.doi.org/10.1175/2008JCLI2690.1>.
- Kelly, K.A., Small, R.J., Samelson, R.M., Qiu, B., Joyce, T.M., Kwon, Y.O., Cronin, M.F., 2010. Western boundary currents and frontal air-sea interaction: gulf stream and Kuroshio extension. *J. Clim.* 23, 5644–5667. <http://dx.doi.org/10.1175/2010JCLI3346.1>.
- Köhl, A., 2015. Evaluation of the GECCO2 ocean synthesis: transports of volume, heat and freshwater in the Atlantic. *Q. J. R. Meteorol. Soc.* 141, 166–181. <http://dx.doi.org/10.1002/qj.2347>.
- Kwon, Y.O., Alexander, M.A., Bond, N.A., Frankignoul, C., Nakamura, H., Qiu, B., Thompson, L., 2010. Role of the gulf stream and Kuroshio–Oyashio systems in large-scale atmosphere-ocean interaction: a review. *J. Clim.* 23, 3249–3281. <http://dx.doi.org/10.1175/2010JCLI3343.1>.
- Kwon, Y.O., Joyce, T.M., 2013. Northern hemisphere winter atmospheric transient eddy heat fluxes and the gulf stream and Kuroshio–Oyashio extension variability. *J. Clim.* 26, 9839–9859. <http://dx.doi.org/10.1175/JCLI-D-12-00647.1>.
- Lee, T., Awaji, T., Balmaseda, M.A., Greiner, E., Stammer, D., 2009. Ocean state estimation for climate research. *Oceanography* 22, 160–167. <http://dx.doi.org/10.5670/oceanog.2009.74>.
- Lozier, M.S., Pratt, L.J., Rogerson, A.M., Miller, P.D., 1997. Exchange geometry revealed by float trajectories in the Gulf stream. *J. Phys. Oceanogr.* 27, 2327–2341. [http://dx.doi.org/10.1175/1520-0485\(1997\)027<2327:EGRFT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1997)027<2327:EGRFT>2.0.CO;2).
- Madec, G., the NEMO team, 2008. NEMO Ocean Engine. Note du Pôle de Modélisation. Institut Pierre-Simon Laplace (IPSL), France No 27, ISSN No 1288–1619.
- Marshall, J., Adcroft, A., Hill, C., Perelman, L., Heisey, C., 1997. A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.* 102, 5753–5766. <http://dx.doi.org/10.1029/96JC02775>.
- McCarthy, G., Frajka-Williams, E., Johns, W.E., Baringer, M.O., Meinen, C.S., Bryden, H.L., Rayner, D., Duchez, A., Roberts, C., Cunningham, S.A., 2012. Observed inter-annual variability of the Atlantic meridional overturning circulation at 26.5°N. *Geophys. Res. Lett.* 39 <http://dx.doi.org/10.1029/2012GL052933>. L19 609.
- McCarthy, G.D., Smeed, D.A., Johns, W.E., Frajka-Williams, E., Moat, B.I., Rayner, D., Baringer, M.O., Meinen, C.S., Collins, J., Bryden, H.L., 2015. Measuring the Atlantic meridional overturning circulation at 26°N. *Progress Oceanogr.* 130, 91–111. <http://dx.doi.org/10.1016/j.pocan.2014.10.006>.
- Meinen, C.S., Baringer, M.O., Garcia, R.F., 2010. Florida current transport variability: an analysis of annual and longer-period signals. *Deep. Res. Part I Oceanogr. Res. Pap.* 57, 835–846. <http://dx.doi.org/10.1016/j.dsr.2010.04.001>.
- Menemenlis, D., Campin, J.-M., Heimbach, P., Hill, C., Lee, T., Nguyen, A., Schodlok, M., Zhang, H., 2008. ECCO2: high resolution global ocean and sea ice data synthesis. *Mercat. Ocean Q. Newsl.* 31, 13–21.
- Menemenlis, D., 2016. Personal communication.
- Miyoshi, T., 2005. Ensemble Kalman filter experiments with a primitive-equation global model. Ph.D. thesis. University of Maryland.
- Mulet, S., Rio, M.H., Mignot, A., Guinehut, S., Morrow, R., 2012. A new estimate of the global 3D geostrophic ocean circulation based on satellite data and in-situ measurements. *Deep. Res. Part II Top. Stud. Oceanogr.* 77–80, 70–81. <http://dx.doi.org/10.1016/j.dsr2.2012.04.012>.
- Nye, J.A., Joyce, T.M., Kwon, Y.-O., Link, J.S., 2011. Silver hake tracks changes in Northwest Atlantic circulation. *Nat. Commun.* 2, 412. <http://dx.doi.org/10.1038/ncomms1420>.
- Palter, J.B., 2015. The role of the Gulf Stream in European climate. *Ann. Rev. Mar. Sci.* 7, 113–137. <http://dx.doi.org/10.1146/annurev-marine-010814-015656>.
- Pacanowski, R.C., Griffies, S.M., 2000. The MOM3 manual. Technical Report. Geophysical Fluid Dynamics Laboratory. http://mdl-mom5.herokuapp.com/web/docs/project/MOM3_manual.pdf (accessed 6.20.17).
- Ratsimandresy, A.W., Pelegri, J.L., 2005. Vertical alignment of the Gulf stream. *Tellus A* 57, 691–700. <http://dx.doi.org/10.1111/j.1600-0870.2005.00115.x>.
- Rossby, T., Flagg, C.N., Donohue, K., Sanchez-Franks, A., Lillibridge, J., 2014. On the long-term stability of Gulf stream transport based on 20 years of direct measurements. *Geophys. Res. Lett.* 41, 114–120. <http://dx.doi.org/10.1002/2013GL058636>.
- Rossby, T., Gottlieb, E., 1998. The oleander project: monitoring the variability of the gulf stream and adjacent waters between New Jersey and Bermuda. *Bull. Am. Meteorol. Soc.* 79, 5–18. [http://dx.doi.org/10.1175/1520-0477\(1998\)079<0005:TOPMTV>2.0.CO;2](http://dx.doi.org/10.1175/1520-0477(1998)079<0005:TOPMTV>2.0.CO;2).
- Saba, V.S., Hyde, K.J.W., Rebeck, N.D., Friedland, K.D., Hare, J.A., Kahru, M., Fogarty, M.J., 2015. Physical associations to spring phytoplankton biomass interannual variability in the U.S. Northeast continental shelf. *J. Geophys. Res. G Biogeosci.* 120, 205–220. <http://dx.doi.org/10.1002/2014JG002770>.
- Saha, S., Moorthi, S., Pan, H.L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.T., Chuang, H.Y., Juang, H.M.H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Van

- Den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R.W., Rutledge, G., Goldberg, M., 2010. The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* 91, 1015–1057. <http://dx.doi.org/10.1175/2010BAMS3001.1>.
- Samelson, R.M., 1992. Fluid exchange across a meandering jet. *J. Phys. Oceanogr.* 22, 431–444. [http://dx.doi.org/10.1175/1520-0485\(1992\)022<0431:FEAAMJ>2.0.CO;2](http://dx.doi.org/10.1175/1520-0485(1992)022<0431:FEAAMJ>2.0.CO;2).
- Sanchez-Franks, A., 2015. Structure and dynamics of the Gulf Stream's interannual migration east of Cape Hatteras. Ph.D. thesis. Stony Brook University.
- Sasaki, H., Nonaka, M., Masumoto, Y., Sasai, Y., Uehara, H., Sakuma, H., 2008. An eddy-resolving hindcast simulation of the quasiglobal ocean from 1950 to 2003 on the Earth Simulator. In: Hamilton, K., Ohfuchi, W. (Eds.), *High Resolution Numerical Modelling of the Atmosphere and Ocean*. Springer New York, New York, NY, pp. 157–185. http://dx.doi.org/10.1007/978-0-387-49791-4_10.
- Schmitz, W.J., Richardson, P.L., 1991. On the sources of the Florida current. *Deep Sea Res. Part A Oceanogr. Res. Pap.* 38, S379–S409. [https://doi.org/10.1016/S0198-0149\(12\)80018-5](https://doi.org/10.1016/S0198-0149(12)80018-5).
- Stammer, D., Awaji, T., Balmaseda, M.A., Behringer, D., Carton, J., Ferry, N., Fischer, A., Fukumori, I., Giese, B., Haines, K., Harrison, E., Heimbach, P., Kamachi, M., Keppenne, C., Masina, S., Menemenlis, D., Ponte, R.M., Remy, E., Rienecker, M.M., Rosati, A., Weaver, A., Wunsch, C.I., Xue, Y., 2010. Ocean information provided through ensemble ocean syntheses. In: Hall, J., Harrison, D.E., Stammer, D. (Eds.), *Proceedings of the OceanObs'09: Sustained Ocean Observations and Information for Society*. Venice, Italy, 21–25 September 2009. Vol. 2 ESA Publication WPP-306. <http://dx.doi.org/10.5270/OceanObs09.cwp.85>.
- Stammer, D., Balmaseda, M., Heimbach, P., Köhl, A., Weaver, A., 2016. Ocean data assimilation in support of climate applications: status and perspectives. *Ann. Rev. Mar. Sci.* 8, 491–518. <http://dx.doi.org/10.1146/annurev-marine-122414-034113>.
- Storto, A., Masina, S., 2016. C-GLORSv5: an improved multipurpose global ocean eddy-permitting physical reanalysis. *Earth Syst. Sci. Data* 8, 679–696. <http://dx.doi.org/10.5194/essd-8-679-2016>.
- Sweet, W., Fett, R., Kerling, J., Violette, La, 1981. Air-sea interaction effects in the lower troposphere across the north wall of the Gulf Stream. *Mon. Weather Rev.* 109, 1042–1052. [http://dx.doi.org/10.1175/1520-0493\(1981\)109<1042:ASIEIT>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1981)109<1042:ASIEIT>2.0.CO;2).
- Szuts, Z.B., Meinen, C.S., 2017. Florida current salinity and salinity transport: mean and decadal changes. *Geophys. Res. Lett.* 44. <http://dx.doi.org/10.1002/2017GL074538>.
- 2017GL074538.
- Taylor, A.H., 1996. North-south shifts of the Gulf stream: ocean-atmosphere interactions in the North Atlantic. *Int. J. Climatol.* 16, 559–583. [http://dx.doi.org/10.1002/\(SICI\)1097-0088\(199605\)16:5<559::AID-JOC26>3.0.CO;2-Z](http://dx.doi.org/10.1002/(SICI)1097-0088(199605)16:5<559::AID-JOC26>3.0.CO;2-Z).
- Taylor, A.H., Stephens, J.A., 1998. The North Atlantic oscillation and the latitude of the Gulf stream. *Tellus A* 50, 134–142. <http://dx.doi.org/10.3402/tellusa.v50i1.14517>.
- Todd, R.E., Owens, W.B., Rudnick, D.L., 2016. Potential vorticity structure in the North Atlantic western boundary current from underwater glider observations. *J. Phys. Oceanogr.* 46, 327–348. <http://dx.doi.org/10.1175/JPO-D-15-0112.1>.
- Toyoda, T., Fujii, Y., Kuragano, T., Kamachi, M., Ishikawa, Y., Masuda, S., Sato, K., Awaji, T., Hernandez, F., Ferry, N., Guinehut, S., Martin, M.J., Peterson, K.A., Good, S.A., Valdivieso, M., Haines, K., Storto, A., Masina, S., Köhl, A., Zuo, H., Balmaseda, M., Yin, Y., Shi, L., Alves, O., Smith, G., Chang, Y.-S., Vernieres, G., Wang, X., Forget, G., Heimbach, P., Wang, O., Fukumori, I., Lee, T., 2017. Intercomparison and validation of the mixed layer depth fields of global ocean syntheses. *Clim. Dyn.* 49, 753–773. <http://dx.doi.org/10.1007/s00382-015-2637-7>.
- Whitaker, J.S., Hamill, T.M., Wei, X., Song, Y., Toth, Z., 2008. Ensemble data assimilation with the NCEP global forecast system. *Mon. Weather Rev.* 136, 463–482. <http://dx.doi.org/10.1175/2007MWR2018.1>.
- Xu, X., Schmitz, W.J., Hurlburt, H.E., Hogan, P.J., 2012. Mean Atlantic meridional overturning circulation across 26.5°N from eddy-resolving simulations compared to observations. *J. Geophys. Res.* 117, C03042. <https://doi.org/10.1029/2011JC007586>.
- Xu, X., Rhines, P.B., Chassignet, E.P., 2016. Temperature–salinity structure of the north atlantic circulation and associated heat and freshwater transports. *J. Clim.* 29, 7723–7742. <https://doi.org/10.1175/JCLI-D-15-0798.1>.
- Yang, S.-C., Corazza, M., Carrassi, A., Kalnay, E., Miyoshi, T., 2009. Comparison of local ensemble transform Kalman filter, 3DVAR, and 4DVAR in a quasigeostrophic model. *Mon. Weather Rev.* 137, 693–709. <http://dx.doi.org/10.1175/2008MWR2396.1>.
- Zhang, R., Vallis, G.K., 2007. The role of bottom vortex stretching on the path of the North Atlantic western boundary current and on the northern recirculation gyre. *J. Phys. Oceanogr.* 37, 2053–2080. <http://dx.doi.org/10.1175/JPO3102.1>.
- Young, G.S., Sikora, T.D., 2003. Mesoscale stratocumulus bands caused by Gulf stream meanders. *Mon. Weather Rev.* 131, 2177–2191. [http://dx.doi.org/10.1175/1520-0493\(2003\)131<2177:MSBCBG>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(2003)131<2177:MSBCBG>2.0.CO;2).
- Zuo, H., Balmaseda, M.A., Mogensen, K., 2015. The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals. *Clim. Dyn.* 1–21. <http://dx.doi.org/10.1007/s00382-015-2675-1>.