On the dynamics of the subtropical mode water from an ensemble view

Luolin Sun[a,b,c], Takaya Uchida[b,c,d], Thierry Penduff[b], William K. Dewar[a,b], Bruno Deremble[b], Andrew C. Poje[e], Eric P. Chassignet[a,c], and Nicolas Wienders[c]

[a] *Department of Earth, Ocean and Atmospheric Science, Florida State University, Florida, USA*

[b] *Université Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP, Institut des Géosciences de l'Environnement, Grenoble, France*

[c] *Center for Ocean-Atmospheric Prediction Studies, Florida State University, Florida, USA*

[d] *Climate Dynamics Laboratory, Center for Earth Sciences, Moscow Institute of Physics and Technology, Dolgoprudny, Russia*

[e] *Department of Mathematics, College of Staten Island, The City University of New York, New York, USA*

Corresponding author: Luolin Sun, ls24bk@fsu.edu

ABSTRACT: Subtropical mode water (STMW) in the North Atlantic has received particular attention due to its formation through direct ocean-atmosphere interactions and its crucial role in upper-ocean ventilation. Here, we diagnose the annual life cycle of STMW from a dynamical perspective and investigate regional mechanisms governing its maintenance. We implement an ensemble of 48 North Atlantic Ocean simulations at mesoscale-permitting resolution, forced by the same atmospheric variability under slightly perturbed initial conditions. This ensemble approach, together with the thickness-weighted averaging (TWA) formalism, enable us to separate the spatiotemporally dependent residual-mean flow from the residual-eddy fluctuations. We characterise STMW as a pool of low potential vorticity (PV) in a buoyancy-coordinate framework within a defined control volume. Our results reveal that the evolution of this low PV pool is primarily governed by the residual-mean PV flux, associated with the residual-mean flow, and partially offset by residual-eddy effects, as quantified by the residual-eddy PV flux. In an integrated and annually averaged sense, the residual-mean PV flux extracts PV from the control volume, contributing to the formation of STMW, whereas the residual-eddy PV flux mixes PV down the gradient, contributing to the erosion of STMW. Diabatic effects become noticeable only when STMW layers are ventilated and exert a comparatively weaker influence on STMW formation. The bolus eddy flux embedded within the residual-mean PV flux becomes pronounced within the control volume of STMW during the formation phase, contributing to PV homogenization by mixing newly formed low PV water with the pre-existing counterpart.

1. Introduction

In the North Atlantic, the subtropical mode water (STMW) is a voluminous water mass characterised by nearly homogeneous properties in the upper ocean. It was initially referred to as the eighteen degree water (EDW) by Worthington (1958) since a large amount of water mass with temperature around 18°C was found in the Sargasso Sea. EDW plays a crucial role in regulating the upper ocean circulation by transporting atmospherically driven properties, such as temperature and carbon dioxide, from the surface to the interior (Hanawa and Talley 2001). It is directly ventilated at the surface in winter, then subducted from the mixed layer into the permanent thermocline in spring and summer. Given by the characteristic structure of the potential vorticity (PV) in the thermocline (Rhines and Young 1982; Luyten et al. 1983; Haynes and McIntyre 1987; Marshall et al. 2001), EDW can be characterised as a weakly stratified water mass with a local minimum PV. Its ventilation can be understood as a consequence of the atmosphere inducing low PV at the outcropping region, which is then mixed with the surrounding higher PV by mesoscale eddies, resulting in a pool of homogeneous, low PV water (Peng et al. 2006).

Various studies have documented that the annual life cycle of EDW is associated with buoyancy loss and mixing processes in the Gulf Stream region where PV gradients are sharper and eddy activity is strong. Maze et al. (2009), Joyce (2013) and Forget et al. (2011) attributed EDW formation and erosion to the intense buoyancy-driven heat loss and gain at the surface south of the Gulf Stream. Billheimer and Talley (2016) found that vertical eddy mixing dominates the erosion of EDW and is more pronounced near Gulf Stream than in the southern subtropical gyre. The PV extraction due to such a buoyancy loss at the surface has been hypothesised to be the source of EDW formation (Maze and Marshall 2011; Maze et al. 2013; Olsina et al. 2013), but was found to only marginally contribute to the total PV flux in Deremble and Dewar (2013). Despite these findings, the evolution of EDW and its relation associated with the Gulf Stream and the subtropical recirculation have not been fully addressed; this will be one of the main focuses of our study.

The role of eddies in EDW formation was initially discussed within a quasi-geostrophic framework by Dewar (1986). By assuming a downgradient eddy PV flux, the maintenance of the low-PV pool was demonstrated by a balance between eddy PV fluxes and buoyant PV forcing. Maze et al. (2013) were the first to use a realistic ocean model to quantify eddy effects, comparing surface PV fluxes computed from an eddy-rich (1/12°) model with those derived from the non-eddy-resolving

OCCA dataset (Forget 2010), and suggested that eddy contributions were secondary. In contrast, Deremble and Dewar (2013), using an eddy-rich ocean model, found that eddies contributed approximately one-quarter of the total PV budget and hypothesized that the maintenance of EDW results from a lateral balance between the mean flow and eddies, consistent with earlier quasi-geostrophic (QG) results (Dewar 1986). In their study, eddies and the mean flow were separated using a thickness-weighted time-averaging operator. Building on these insights, we are motivated to consider a time-dependent mean-eddy separation to explicitly study the temporal effects of the mean flow and eddies on the annual cycle of STMW. In addition, we will investigate the role of eddies in subducting surface-formed EDW into the main thermocline.

In this study, we quantify the North Atlantic STMW by identifying it as a low potential vorticity pool, and hence refer to it as STMW rather than EDW, as our definition is based on PV structure rather than temperature homogeneity. To examine the respective roles of the mean flow and eddies, we employ an ensemble of eddy-permitting North Atlantic simulations. The mean flow is defined as the spatiotemporally varying ensemble-mean variable, enabling us to separate eddy fields without prescribing assumptions on their temporal or spatial scales.

The ensemble-averaging approach adopted in this study naturally leads to the use of the thickness-weighted averaging (TWA) framework (Young 2012), in which dynamical variables are both thickness-weighted and ensemble-averaged within buoyancy layers. This TWA framework yields a PV conservation analogous to that of the Ertel's potential vorticity, traditionally used to study the PV budget associated with EDW formation and evolution (Marshall et al. 2001; Czaja and Hausmann 2009; Maze and Marshall 2011; Olsina et al. 2013; Deremble and Dewar 2013). Water-mass formation (Walin 1982), the primary factor of STMW formation (Joyce et al. 2013; Maze and Marshall 2011), is implicitly embedded in the TWA formalism, and is naturally connected to the TWA's PV conservation. Specifically, the thickening of water masses due to surface buoyancy variations is reflected as an increase in layer thickness, which, given thickness-weighted PV conservation within the layer, leads to a decrease in the TWA's PV. The framework implemented in this study captures the key characteristics of the STMW annual life cycle, linking the seasonal thickening and thinning of layers directly to the evolution of PV.

The article is structured as follows. In Section 2, we introduce the ensemble simulations used in this study and briefly review the TWA formalism. In Section 3, we characterize the simulated

STMW and describe its annual life cycle. Later in the section, we define a control volume for a low PV pool within a STMW layer. This is then followed by a diagnosis of the evolution of the PV budget within this control volume in Section 4, with the results interpreted in terms of the annual characteristics of STMW formation and erosion. In Section 5, we investigate the regional mechanisms governing STMW maintenance within the subtropical gyre. In Section 6, we discuss the role of the eddy-induced bolus velocity on the lateral PV mixing during the formation phase of STMW. We discuss technical challenges related to closing the PV budget in Section 7, and summarise the main findings of the study in Section 8.

2. Methodology

We begin this section by describing the ensemble simulations used in the study and introducing the ensemble-based separation of mean and eddy components. We then revisit the TWA framework introduced by Young (2012) and derive the corresponding PV equation. Mathematical notations are summarized in Table 1.

a. Ensemble simulations

We employ an ensemble of North Atlantic simulations (NA12), spanning 20°S to 55°N in latitude as described in Jamet et al. (2019, 2020) and Uchida et al. (2025), using the Massachusetts Institute of Technology general circulation model (MITgcm; Marshall et al. 1997). These simulations are mesoscale-permitting at 1/12° and partially coupled to the atmosphere via the Cheap Atmospheric Mixed Layer model (Deremble et al. 2013). Forty-eight ensemble members are initialized with distinct but physically consistent oceanic states, exposed to the same prescribed atmospheric variability beginning on January 1st, 1963 and then simulated onwards for five years. Instantaneous snapshots of the model state are saved every five days during the last year, 1967, and are the model outputs used in this study.

We follow the prescription described in Uchida et al. (2022, 2023) to interpolate the variable outputs from geopotential coordinates (t,x,y,z) into buoyancy coordinates represented by $(\tilde{t},\tilde{x},\tilde{y},\tilde{b})$. Here, $\tilde{b}$ is the thermodynamically and dynamically consistent buoyancy; it is defined via in-situ density anomaly (see also Stanley and Marshall 2022), and is close to being a neutral surface and has a monotonic relation with depth z. $\tilde{x}$ and $\tilde{y}$ are the lateral coordinates on the $\tilde{b}$-surfaces. We assign the same $\tilde{b}$-coordinate system to all ensemble members. Consequently, the corresponding depths are considered to be a function of $\tilde{b}$. For each ensemble member, indexed by n, the depth z is transformed from geopotential coordinates to $\tilde{b}$-coordinates as

$$z_n = \zeta_n(\tilde{t}, \tilde{x}, \tilde{y}, \tilde{b}), \quad n = 1, \ldots, N, \tag{1}$$

and the thickness of a buoyancy layer centered at $\tilde{b}$ is defined as

$$\sigma_n(\tilde{t}, \tilde{x}, \tilde{y}, \tilde{b}) \stackrel{\text{def}}{=} \frac{\partial \zeta_n}{\partial \tilde{b}}, \quad n = 1, \ldots, N, \tag{2}$$

where $N = 48$ is the total number of ensemble members.

The ensemble-based Reynolds' decomposition is well suited for the present study of the evolution of STMW, as it enables a separation between mean variables and eddy fields at any location and time without making implicit assumptions about their spatial or temporal scales. A full velocity field $\mathbf{u}_n$ is decomposed into an ensemble-mean field

$$\langle \mathbf{u} \rangle \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^{N} \mathbf{u}_n, \tag{3}$$

where $\langle \cdot \rangle$ denotes the ensemble-averaging operator, and an ensemble-eddy field

$$\mathbf{u}'_n \stackrel{\text{def}}{=} \mathbf{u}_n - \langle \mathbf{u} \rangle. \tag{4}$$

$\langle \mathbf{u} \rangle$ can be considered as the common oceanic response among members to the same prescribed atmospheric variability, and $\mathbf{u}'_n$ as the ocean-generated intrinsic eddies (Penduff et al. 2011; Leroux et al. 2018; Constantinou and Hogg 2021). Similarly, the ensemble-mean depth and layer thickness are given by

$$\langle \zeta \rangle \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^{N} \zeta_n, \tag{5a}$$

$$\langle \sigma \rangle \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=1}^{N} \sigma_n, \tag{5b}$$

and their eddy fluctuations around their ensemble-mean fields are given by

$$\zeta_n' \stackrel{\text{def}}{=} \zeta_n - \langle\zeta\rangle, \tag{6a}$$

$$\sigma_n' \stackrel{\text{def}}{=} \sigma_n - \langle\sigma\rangle. \tag{6b}$$

$\langle\zeta\rangle$ and $\langle\sigma\rangle$ characterise the common features of the stratifications among the ensemble members, and ζ_n' and σ_n' are displacements around them. These variables play a key role in the TWA framework as well as in the analysis of STMW dynamics. Before presenting details of the TWA formalism, we remind readers that all calculations of the variables and their analyses throughout this study are carried out in the $\tilde{b}$-coordinate system.

b. Governing equations

1) Momentum equations

In TWA, the primitive equations are reformulated as the following lateral residual-mean momentum equations and the ensemble-mean thickness equation:

$$\hat{u}_{\tilde{t}} + \hat{u}\hat{u}_{\tilde{x}} + \hat{v}\hat{u}_{\tilde{y}} + \hat{\varpi}\hat{u}_{\tilde{b}} - f\hat{v} + \langle m\rangle_{\tilde{x}} + \boldsymbol{\nabla}\cdot\mathbf{E}^u = \hat{\mathcal{X}}, \tag{7a}$$

$$\hat{v}_{\tilde{t}} + \hat{u}\hat{v}_{\tilde{x}} + \hat{v}\hat{v}_{\tilde{y}} + \hat{\varpi}\hat{v}_{\tilde{b}} + f\hat{u} + \langle m\rangle_{\tilde{y}} + \boldsymbol{\nabla}\cdot\mathbf{E}^v = \hat{\mathcal{Y}}, \tag{7b}$$

$$\langle\sigma\rangle_{\tilde{t}} + \langle\sigma\rangle\boldsymbol{\nabla}\cdot\hat{\mathbf{u}} = 0,[1] \tag{7c}$$

where the subscripts denote the partial differentiations in $\tilde{b}$-coordinates and $\widehat{(\cdot)}$ denotes a TWA operator. $\hat{u}$, $\hat{v}$ and $\hat{\varpi}$ are residual-mean velocities of the form

$$\hat{\theta} \stackrel{\text{def}}{=} \frac{\langle\sigma\theta\rangle}{\langle\sigma\rangle}, \quad \theta \in \{u, v, \varpi\}. \tag{8}$$

$\hat{\mathbf{u}} \stackrel{\text{def}}{=} (\hat{u}, \hat{v}, \hat{\varpi})$ is referred to as the TWA velocity, and its divergence governs the per unit rate of change in thickness

$$-\langle\sigma\rangle^{-1}\langle\sigma\rangle_{\tilde{t}} = \boldsymbol{\nabla}\cdot\hat{\mathbf{u}} \neq 0. \tag{9}$$

which is generally non-zero for a non-stationary flow.

[1] We rewrite Eq. (68) in Young (2012) by using the vector notation (12).

The eddy forcing is given by the divergences of the three-dimensional Eliassen-Palm (E-P) flux vectors,

$$\mathbf{E}^u \stackrel{\text{def}}{=} \left(\widehat{u''u''} + \frac{\langle \zeta'^2 \rangle}{2\langle \sigma \rangle}\right)\langle \mathbf{e} \rangle_1 + \widehat{v''u''}\langle \mathbf{e} \rangle_2 + \left(\widehat{\varpi''u''} + \frac{\langle \zeta' m'_{\tilde{x}} \rangle}{\langle \sigma \rangle}\right)\langle \mathbf{e} \rangle_3 \,, \tag{10a}$$

$$\mathbf{E}^v \stackrel{\text{def}}{=} \widehat{u''v''}\langle \mathbf{e} \rangle_1 + \left(\widehat{v''v''} + \frac{\langle \zeta'^2 \rangle}{2\langle \sigma \rangle}\right)\langle \mathbf{e} \rangle_2 + \left(\widehat{\varpi''v''} + \frac{\langle \zeta' m'_{\tilde{y}} \rangle}{\langle \sigma \rangle}\right)\langle \mathbf{e} \rangle_3 \,. \tag{10b}$$

$\mathbf{u}''_n \stackrel{\text{def}}{=} (u''_n, v''_n, \varpi''_n)$ is the residual-eddy velocity for each ensemble member with the components

$$(u''_n, v''_n, \varpi''_n) \stackrel{\text{def}}{=} (u_n, v_n, \varpi_n) - (\hat{u}, \hat{v}, \hat{\varpi}) \,, \tag{11}$$

and modifies the corresponding thickness fluctuation σ'_n. A schematic of the coordinate transformation and the decomposition between the residual-mean and the residual-eddy velocities is shown in Figure 1.

The divergence operator is in the form

$$\boldsymbol{\nabla} \cdot \left(E^1 \langle \mathbf{e} \rangle_1 + E^2 \langle \mathbf{e} \rangle_2 + E^3 \langle \mathbf{e} \rangle_3\right) \stackrel{\text{def}}{=} \langle \sigma \rangle^{-1} \left(\langle \sigma \rangle E^1\right)_{\tilde{x}} + \langle \sigma \rangle^{-1} \left(\langle \sigma \rangle E^2\right)_{\tilde{y}} + \langle \sigma \rangle^{-1} \left(\langle \sigma \rangle E^3\right)_{\tilde{b}} \,, \tag{12}$$

where $\langle \mathbf{e} \rangle_1 = \mathbf{i} + \langle \zeta \rangle_{\tilde{x}} \mathbf{k}$, $\langle \mathbf{e} \rangle_2 = \mathbf{j} + \langle \zeta \rangle_{\tilde{y}} \mathbf{k}$ and $\langle \mathbf{e} \rangle_3 = \langle \sigma \rangle \mathbf{k}$ are the vector basis of the $\tilde{b}$-coordinate system.

The Montgomery potential is related to the dynamically active part of hydrostatic pressure, ϕ, via

$$m(\tilde{t}, \tilde{x}, \tilde{y}, \tilde{b}) \stackrel{\text{def}}{=} \phi(t, x, y, \zeta) - \tilde{b}\zeta(\tilde{t}, \tilde{x}, \tilde{y}, \tilde{b}) \,; \tag{13}$$

and $\hat{\mathcal{X}}$ and $\hat{\mathcal{Y}}$ are the adiabatic terms.

2) Potential vorticity equation

Subtracting the cross derivatives of (7a) and (7b) gives the definition of the PV in TWA formalism

$$\Pi^{\#} \stackrel{\text{def}}{=} \frac{f + \hat{v}_{\tilde{x}} - \hat{u}_{\tilde{y}}}{\langle \sigma \rangle} \,, \tag{14}$$

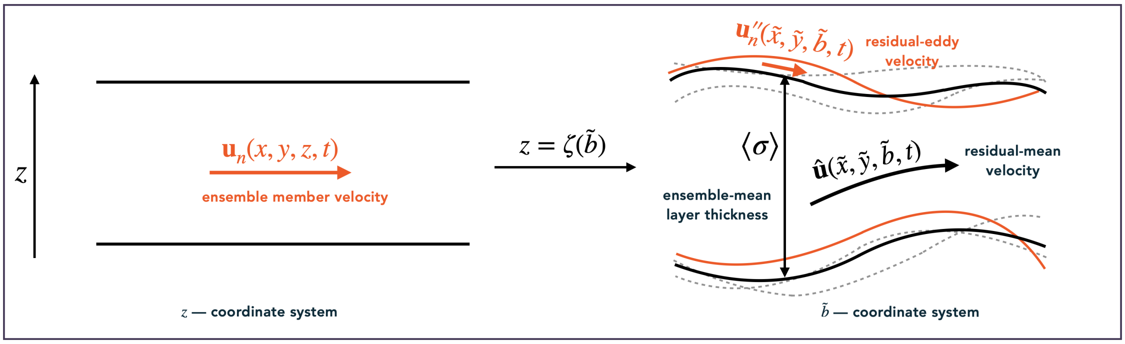


FIG. 1. Conceptual schematic illustration of the TWA formalism. The velocity field $\mathbf{u}_n$ of an ensemble member in geopotential coordinates (left panel) is transformed into buoyancy coordinates, where it is separated into a residual-mean velocity $\hat{\mathbf{u}}$ and a residual-eddy velocity $\mathbf{u}''_n$ (right panel). Quantities and curves shown in orange are specific to the individual ensemble member, while those in black represent ensemble-mean properties shared across members. The vertical structure of the buoyancy coordinates (black curves) is determined by the ensemble-mean layer thickness $\langle\sigma\rangle$. Dashed lines denote layers from other ensemble members.

and leads to a conservation equation for PV of $\hat{\mathbf{u}}$ and $\langle\sigma\rangle$,

$$\langle\sigma\rangle^{-1}\left(\langle\sigma\rangle\Pi^{\#}\right)_{\tilde{t}} + \nabla\cdot\mathbf{J}^{\#} + \nabla\cdot\mathbf{F}^{\#} + \nabla\cdot\mathbf{G}^{\#} = 0\,. \tag{15}$$

(15) is analogous to the conservation form of the Ertel's PV discussed in Haynes and McIntyre (1987); Marshall et al. (2001). The impermeability theorem holds here so $\Pi^{\#}$ of the buoyancy layer can only be modified at the ocean boundaries. We remind readers that $\Pi^{\#}$ is not a mean quantity of PV but rather the PV associated with the residual-mean flow $\hat{\mathbf{u}}$ (viz. $\Pi^{\#} \neq \widehat{\Pi}$; Young 2012), as the TWA operator does not commute with cross-product (Maddison and Marshall 2013). Throughout the following discussion, we refer to $\Pi^{\#}$ simply as 'PV' for brevity.

The residual-mean PV flux

$$\mathbf{J}^{\#} \overset{\text{def}}{=} \hat{u}\Pi^{\#}\langle\mathbf{e}\rangle_1 + \hat{v}\Pi^{\#}\langle\mathbf{e}\rangle_2 + \hat{\varpi}\Pi^{\#}\langle\mathbf{e}\rangle_3\,, \tag{16}$$

quantifies the advection of $\Pi^{\#}$ by the TWA velocity $\hat{\mathbf{u}}$. It consists both the large-scale advection by the ensemble-mean velocity $\langle\mathbf{u}\rangle$ and the lateral eddy-induced advection by the bolus velocity

$$\mathbf{u}^{*} \overset{\text{def}}{=} \frac{\langle\sigma'\mathbf{u}'_h\rangle}{\langle\sigma\rangle} = \hat{\mathbf{u}} - \langle\mathbf{u}\rangle\,, \tag{17}$$

where $\mathbf{u}'_h \stackrel{\text{def}}{=} u'\langle\mathbf{e}\rangle_1 + v'\langle\mathbf{e}\rangle_2$ is along $\tilde{b}$-surfaces. In TWA, $\mathbf{u}^*$ is designed to be embedded in $\hat{\mathbf{u}}_h$ because variations of buoyancy-driven layers should naturally be incorporated into mean (TWA) variables.

The residual-eddy PV flux

$$\mathbf{F}^\# \stackrel{\text{def}}{=} \left(\langle\sigma\rangle^{-1}\boldsymbol{\nabla}\cdot\mathbf{E}^v\right)\langle\mathbf{e}\rangle_1 - \left(\langle\sigma\rangle^{-1}\boldsymbol{\nabla}\cdot\mathbf{E}^u\right)\langle\mathbf{e}\rangle_2\,, \tag{18}$$

quantifies the effects of the eddy forcing appeared in the momentum equations (7) on the PV transport. The diabatic PV flux is given by

$$\mathbf{G}^\# \stackrel{\text{def}}{=} \langle\sigma\rangle^{-1}\left(-\hat{\mathcal{Y}} + \hat{v}_{\tilde{b}}\hat{\varpi}\right)\langle\mathbf{e}\rangle_1 + \langle\sigma\rangle^{-1}\left(\hat{\mathcal{X}} - \hat{u}_{\tilde{b}}\hat{\varpi}\right)\langle\mathbf{e}\rangle_2 - \hat{\varpi}\Pi^\#\langle\mathbf{e}\rangle_3\,. \tag{19}$$

In the STMW analysis the low PV pool is conventionally determined by the potential vorticity $\Pi^\#$, rather than the absolute vorticity $\langle\sigma\rangle\Pi^\#$ in the tendency term in (15). By applying the chain rule to $(\langle\sigma\rangle\Pi^\#)_{\tilde{t}}$ and using the thickness equation in the form (9), we rewrite the PV equation (15) in an equivalent form

$$\Pi^\#_t - \Pi^\#\boldsymbol{\nabla}\cdot\hat{\mathbf{u}} + \boldsymbol{\nabla}\cdot\mathbf{J}^\# + \boldsymbol{\nabla}\cdot\mathbf{F}^\# + \boldsymbol{\nabla}\cdot\mathbf{G}^\# = 0\,, \tag{20}$$

and use it to study the PV budget. Apart from the flux divergences, the PV tendency is also linked to the $\Pi^\#$-weighted divergence of the TWA velocity and, through (9), to the rate of change of $\langle\sigma\rangle$. This relation indicates how PV and layer thickness in STWM (and so the volume), co-vary during the annual cycle. A detailed derivation of (20) is provided in Appendix A.

If the flow is steady and adiabatic, $\hat{\mathbf{u}}$ aligns with the non-divergent residual velocity (Young 2012; Aoki 2014)

$$\mathbf{u}^\# \stackrel{\text{def}}{=} \hat{u}\langle\mathbf{e}\rangle_1 + \hat{v}\langle\mathbf{e}\rangle_2 + \langle\sigma\rangle^{-1}\left(\langle\zeta_{\tilde{t}}\rangle + \hat{\varpi}\langle\zeta\rangle_{\tilde{b}}\right)\langle\mathbf{e}\rangle_3\,. \tag{21}$$

We do not use $\mathbf{u}^\#$ in the formalism here because it does not naturally appear in the momentum equations (7), and therefore not in any PV conservation form considered in this study. Replacing $\hat{\mathbf{u}}$ with $\mathbf{u}^\#$ would obscure the diabatic effects through the dominant term $\langle\zeta\rangle_{\tilde{t}}$, which does not describe the dynamics of STMW within a buoyancy layer but the layer depth. Moreover, $\langle\zeta\rangle_{\tilde{t}}$ varies strongly during the STMW formation phase and cannot be computed accurately from the 5-day snapshots of offline model outputs used in this study.

Table 1. Summary of key notations and variables used in the study

Mathematical symbol	Description
$\tilde{b}$	Thermodynamically and dynamically consistent buoyancy
z	Depth in geopotential coordinates
ζ	Depth in $\tilde{b}$-coordinates
$\sigma \overset{\text{def}}{=} \zeta_{\tilde{b}}$	Thickness of buoyancy layers
$\langle \cdot \rangle$	Ensemble-averaging operator
$\widehat{(\cdot)}$	Thickness-weighted ensemble-averaging (TWA) operator
$\overline{(\cdot)}$	Time-averaging operator
$\langle \sigma \rangle$	Ensemble-mean thickness
$\hat{u}$	Residual-mean velocity component
$\mathbf{u} \overset{\text{def}}{=} (u, v, \varpi)$	Three-dimensional velocity field in the $\tilde{b}$-coordinate convective derivative
$\langle \mathbf{u} \rangle \overset{\text{def}}{=} (\langle u \rangle, \langle v \rangle, \langle \varpi \rangle)$	Three-dimensional ensemble-mean velocity
$\mathbf{u}' \overset{\text{def}}{=} \mathbf{u} - \langle \mathbf{u} \rangle$	Three-dimensional ensemble-eddy velocity
$\hat{\mathbf{u}} \overset{\text{def}}{=} (\hat{u}, \hat{v}, \hat{\varpi})$	Three-dimensional TWA velocity
$\mathbf{u}'' \overset{\text{def}}{=} \mathbf{u} - \hat{\mathbf{u}}$	Three-dimensional residual-eddy velocity
$\hat{\mathbf{u}}_h \overset{\text{def}}{=} (\hat{u}, \hat{v})$	Lateral TWA velocity on $\tilde{b}$-surfaces
$\mathbf{u}^* \overset{\text{def}}{=} \frac{\langle \sigma' \mathbf{u}'_h \rangle}{\langle \sigma \rangle} = \hat{\mathbf{u}} - \langle \mathbf{u} \rangle$	Two-dimensional bolus-eddy velocity
$\mathbf{E}^u, \mathbf{E}^v$	Eliassen-Palm flux vectors
$(\cdot)^\#$	Non-TWA variables consisting the residual field
$\Pi^\#$	Potential vorticity on a buoyancy layer with ensemble-mean thickness $\langle \sigma \rangle$
$\mathbf{J}^\#$	Residual-mean PV flux
$\mathbf{F}^\#$	Residual-eddy PV flux
$\mathbf{G}^\#$	Diabatic PV flux
$\mathbf{u}^* \Pi^\#$	Bolus-eddy PV flux
$\langle \mathbf{u} \rangle \Pi^\#$	Ensemble-mean PV flux
Ω	Control volume of PV pool in the core STMW layer

3. Characterising STMW in TWA

In this section, we characterise STMW layers in the TWA framework and outline their annual cycle. We then determine a control volume of a low PV pool in the core layer of STMW. The dynamics of PV evolution within this control volume will be the focus of the study.

a. Model evaluation

We evaluate the NA12 simulation over year 1967 by comparing the monthly-mean temperature profiles averaged over all ensemble members with the observation-based Met Office EN4 dataset of $1^\circ \times 1^\circ$ resolution (Good et al. 2013; Cheng et al. 2014). Figure 2 compares NA12 and

EN4 in terms of sea surface temperatures (SST) in February and vertical sections of February and September potential temperatures zonally averaged between 75°W and 50°W. Overall, the ensemble-mean temperature $\langle T \rangle$ from NA12 exhibits a warm bias of up to 3°C and a more pronounced stratification compared to observations. The temperature of the simulated STMW varies between 20°C and 22°C, and accordingly, the winter outcrop is found about 2° south of its observed latitude (see Figures 2a and 2d).

Despite these differences, the vertical and latitudinal structures of stratification are similar between EN4 and NA12. In both datasets the isopycnals where the STMW resides outcrop between 30°N and 40°N in winter (Figures 2b and 2e), and the water masses are isolated in summer (Figures 2c and 2f). Hence, we argue that despite its warm bias, the model dynamics that control the life cycle of the STMW are relevant for the real-world setting. These ensemble simulations were also shown to be highly correlated with the RAPID-MOCHA observations of AMOC (Jamet et al. 2019) and capture the temporal variability along the Deep Western Boundary Current (Uchida et al. 2024), further supporting the dynamical relevance of the simulation in this region.

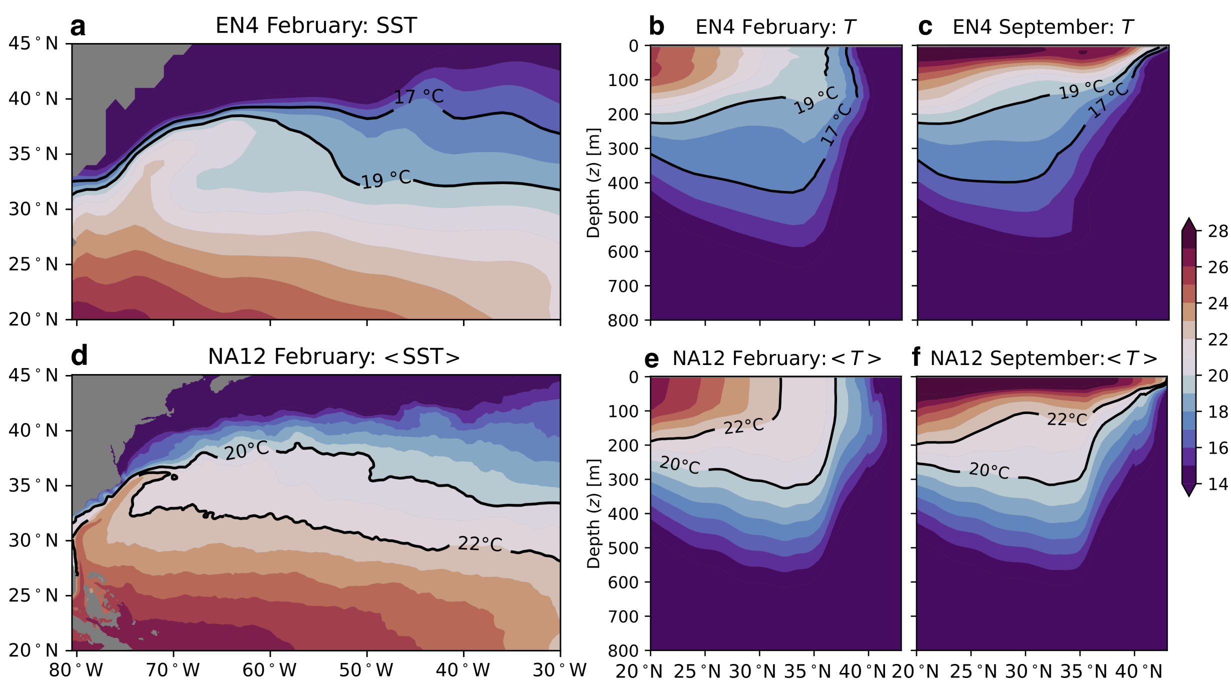


Fig. 2. Sea surface temperature in February 1967 for (a) EN4 and (d) NA12. Monthly temperatures zonally averaged between 75°W and 50°W for EN4 (b,c) and NA12 (e,f). Black contours show the temperature range of the STMW.

b. Simulated STMW layers

Since STMW is formed when the mixed layer deepens in winter, determining the mixed layer depth (MLD) is crucial for understanding the formation mechanism. We determine the MLD for the ensemble as follows. We first estimate the MLD for each ensemble member as a time-varying surface in $\tilde{b}$-coordinates, where the potential density becomes 0.03 kg m^{-3} larger than that at 10 m depth (de Boyer Montégut et al. 2004). Then, to mitigate the influence of local extrema in individual MLD estimates across the ensemble, we define the ensemble MLD as the third quantile of the 48 member estimates[2]. The mixed layer presents a subtlety for the TWA analysis since vertical displacements cannot be perfectly resolved by the buoyancy within the mixed layer. That is, as the mixed layer intrudes into the STMW region, the numerical accuracy of the STMW estimation degrades. In this study, we choose to retain the STMW within the mixed layer, and the rationale for this choice will be explained later in the section.

We identify the simulated STMW by its associated local PV minimum on $\tilde{b}$-surfaces. Figure 3 shows the zonally averaged structure of PV on $\tilde{b}$-surfaces for each month. The two pink lines identify the top and bottom surfaces of the STMW, and correspond to the isolines of the temperatures in Figure 2. The PV is minimum along the surface $\tilde{b} = -0.25$ m s^{-2} which identifies the center of the core layer, shown as red lines in Figure 3. The buoyancy $\tilde{b}$ of the layers within which the STMW sits varies between -0.253 m s^{-2} and -0.247 m s^{-2}. Between February and April, STMW layers slant greatly starting at 30°N, intrude into the deepening mixed layer (its depth is colored in grey) and are directly ventilated at the ocean surface near 35°N. They are then subducted towards the ocean interior from May as the mixed layer shallows. As a consequence of this process, the total thickness of the STMW peaks in June and slowly decreases afterward, in accordance with the formation and erosion of the STMW. Overall, these results demonstrate that the annual life cycle of the STMW is adequately characterized by the PV defined in the thickness-weighted ensemble-averaging framework.

[2] As the MLDs in some members are deeper than others, we have opted for the third quantile instead of the ensemble-mean to ensure that the layers beneath the mixed layer remain substantial within the ensemble. This approach is conceptually similar to using the ensemble mean under a Gaussian distribution, where the second quantile is simply the mean.

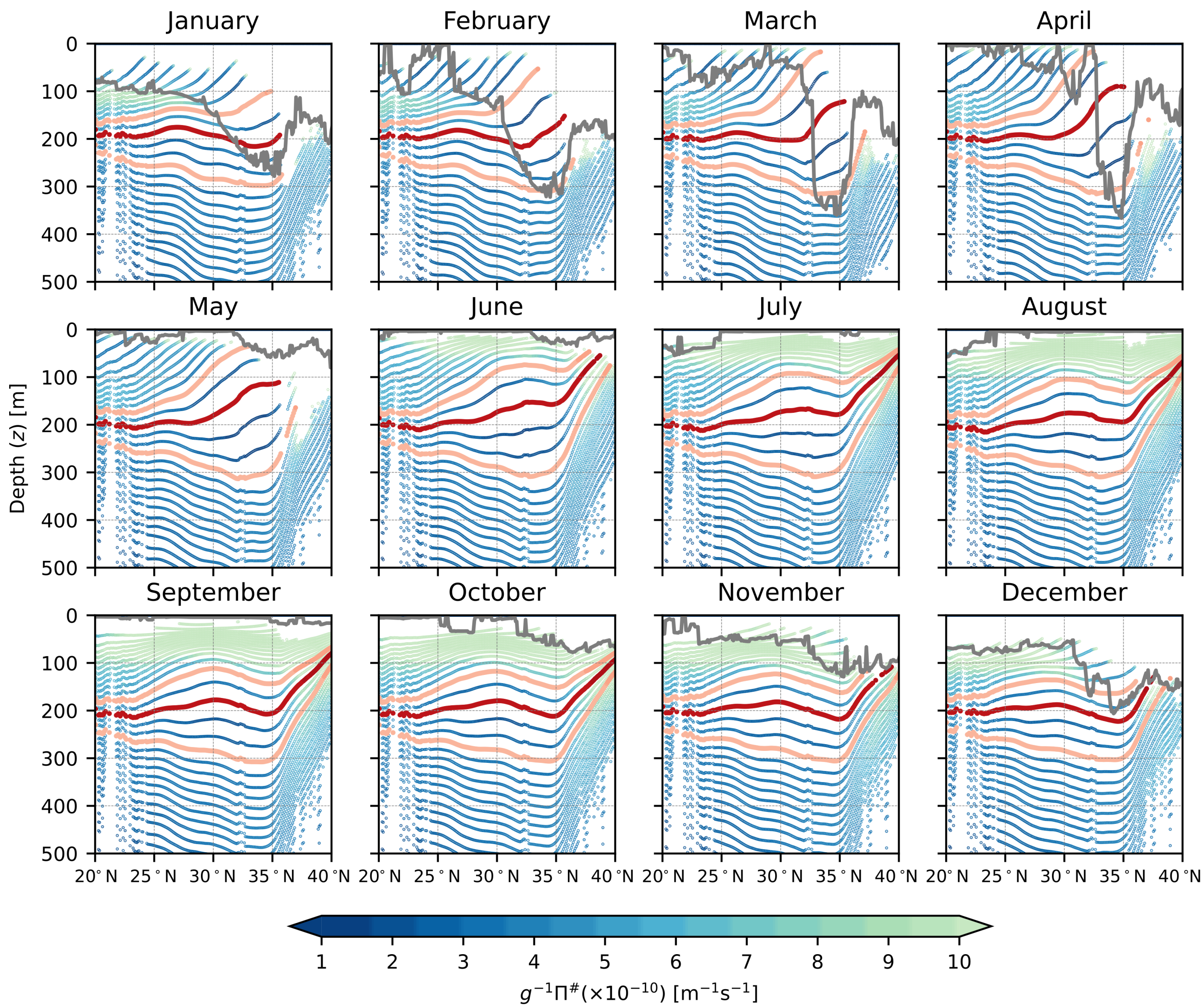


Fig. 3. $\Pi^{\#}$ is divided by g to adopt the conventional unit [m^{-1}s^{-1}] for PV. Blue-green lines show the monthly-averaged PV structure along $\tilde{b}$-surfaces, zonally averaged between 75°W and 50°W. This is the same region chosen for averaging the potential temperatures in Figure 2. Pink lines indicate the centers of the top and bottom layers of the STMW, and the red lines the centers of the core layer $\tilde{b} = -0.25\,\mathrm{m\,s^{-2}}$. The bottom of the zonal mean of the third quantile of MLD is plotted in gray. We remind the reader that the MLD is estimated locally in space-time and then zonally and monthly averaged when plotting. The tips of the lines near the surface show where the layers reaches outcrop region.

c. Control volume of low PV pool

To study the dynamics of the STMW, we focus our study on the core layer. Because it is rarely contaminated by outcropping, where PV becomes ill-defined in the layers above, the analysis can be done for all months of the year. It is also the most representative layer among STMW layers, as its thickness varies the most in time.

The annual-mean ensemble-mean thickness $\overline{\langle\sigma\rangle}$ of the core layer and the annual-mean PV, $\overline{g^{-1}\Pi^{\#}}$, are shown on the center surface in Figure 4. $\overline{(\cdot)}$ denotes the temporal averaging and $\Pi^{\#}$ is divided by g to adopt the conventional unit [m^{-1}s^{-1}] for PV. We restrict our analysis over a region in the core layer that does not outcrop over the year and subsequently exclude the area shaded in light gray in Figure 4b from the analysis. We determine the PV threshold of the STMW as $g^{-1}\Pi^{\#} \leq 3 \times 10^{-10}\,\mathrm{m}^{-1}\,\mathrm{s}^{-1}$. This threshold is larger than those used in some other studies (Hanawa and Talley 2001; Kwon and Riser 2004; Forget et al. 2011; Maze et al. 2013; Billheimer and Talley 2013; Li et al. 2022) because PV is more stratified in our study[3].

Next, we determine a control volume for a low PV pool in this core STMW layer. Although we define STMW using a constant PV threshold, tracking a time-evolving PV contour is technically difficult, especially when $\tilde{b}$-surfaces outcrop at the surface and both mechanical and buoyancy forcing exist. We therefore adopt a fixed lateral boundary C for the control volume, which does not evolve with the instantaneous STMW region. Specifically, C is defined as the contour of the annual-mean field $\overline{g^{-1}\Pi^{\#}} = 3 \times 10^{-10}\,\mathrm{m}^{-1}\,\mathrm{s}^{-1}$, shown as the black contour in both panels. Together with the time-varying height $\mathrm{d}\langle\zeta\rangle = \langle\sigma\rangle\mathrm{d}\tilde{b}$, this control volume Ω encloses the low-PV water in the core layer.

Figure 5 illustrates the monthly evolution of STMW ($g^{-1}\Pi^{\#} \leq 3 \times 10^{-10}\,\mathrm{m}^{-1}\mathrm{s}^{-1}$, colored in dark blue) in the core layer on $\tilde{b} = -0.25\,\mathrm{m\,s^{-2}}$. A portion of STMW persists year-round along the coast between 25°N and 35°N, 80°W and 70°W, and in the southern Sargasso Sea between 25°N and 30°N, 70°W and 60°W. In January, low PV water induced by surface buoyancy loss (Maze and Marshall 2011; Maze et al. 2013) begins to emerge northeast of the control volume and occupies the 'formation zone' by February, as outlined by the green box. This newly formed low PV water then migrates southward and connects with the pre-existing low PV water. The PV minimum during this phase is located near 35°N, 60°W, consistent with the region where Warren (1972) identified the winter formation of STMW. Between March and May, the STMW expands as more low PV water is injected into the system, reaching its maximum extent in June. From July to August, PV within the STMW gradually increases as it mixes with surrounding high PV waters. Beginning

[3]The difference is also potentially attributable to using $\Pi^{\#}$ rather than PV of the Eulerian-mean flow (e.g. $\langle u\rangle, \langle v\rangle, \langle T\rangle$), avoiding potential density in defining PV and/or to our PV is not being normalized by density.

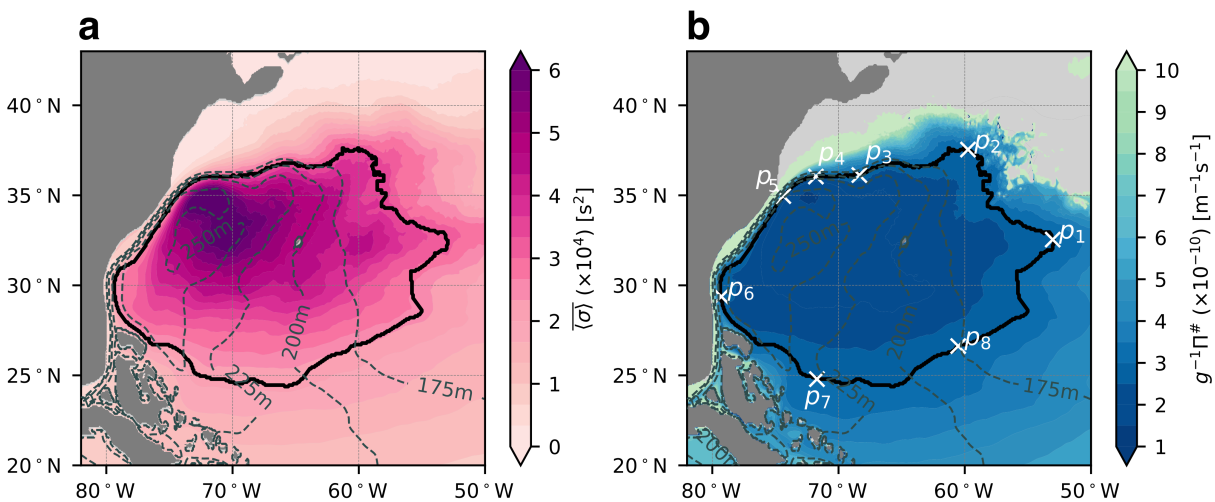


FIG. 4. Annual mean of (a) layer thickness $\overline{\langle\sigma\rangle}$ and (b) PV $\overline{g^{-1}\Pi^{\#}}$ on surface $\tilde{b} = -0.25\,\mathrm{m\,s^{-2}}$ centered in the core STMW layer. Black contours are the boundaries of STMW with the threshold $g^{-1}\Pi^{\#} = 3\times 10^{-10}\,\mathrm{m^{-1}\,s^{-1}}$. Grey dashed lines label the depth of this $\tilde{b}$-surface. p-indices in white are critical locations along the contour, and are labelled by numbers in a counter-clockwise direction following the integration along the contour. Segment from p_7 to p_5 traces the Florida Current along the coastline and p_5 is near Cape Hatteras. Segment from p_5 to p_2 is on the southern flank of the Gulf stream. The separation point where the Gulf Stream detaches from the coast and migrates eastward is between p_5 and p_4. Low PV water formed in early winter enters the control volume from the north through segment between p_2 and p_1. Segment p_1 to p_7 approximately follows the recirculation of the subtropical gyre.

in September, STMW undergoes noticeable erosion, starting in the 'formation zone' as the layer subducts, and continues to shrink, reaching its minimum extent by December, setting the stage for renewal in the following year.

Under stationarity, a closed contour of the Bernoulli potential would be a viable choice for the lateral boundary, since the net PV flux through it vanishes (Marshall 2000; Polton and Marshall 2003, 2007). Such a boundary also seems more suitable for analyzing PV flux dynamics, because the time-mean Bernoulli potential can be regarded as the streamfunction of the time-mean net PV flux when the averaging period is sufficiently long (cf. Appendix B). In our case, however, no suitable contour of the Bernoulli potential could be identified to serve as the lateral boundary of the control volume. Many contours either intersect the western boundary or fail to enclose the region where STMW forms and erodes (see Figure B1).

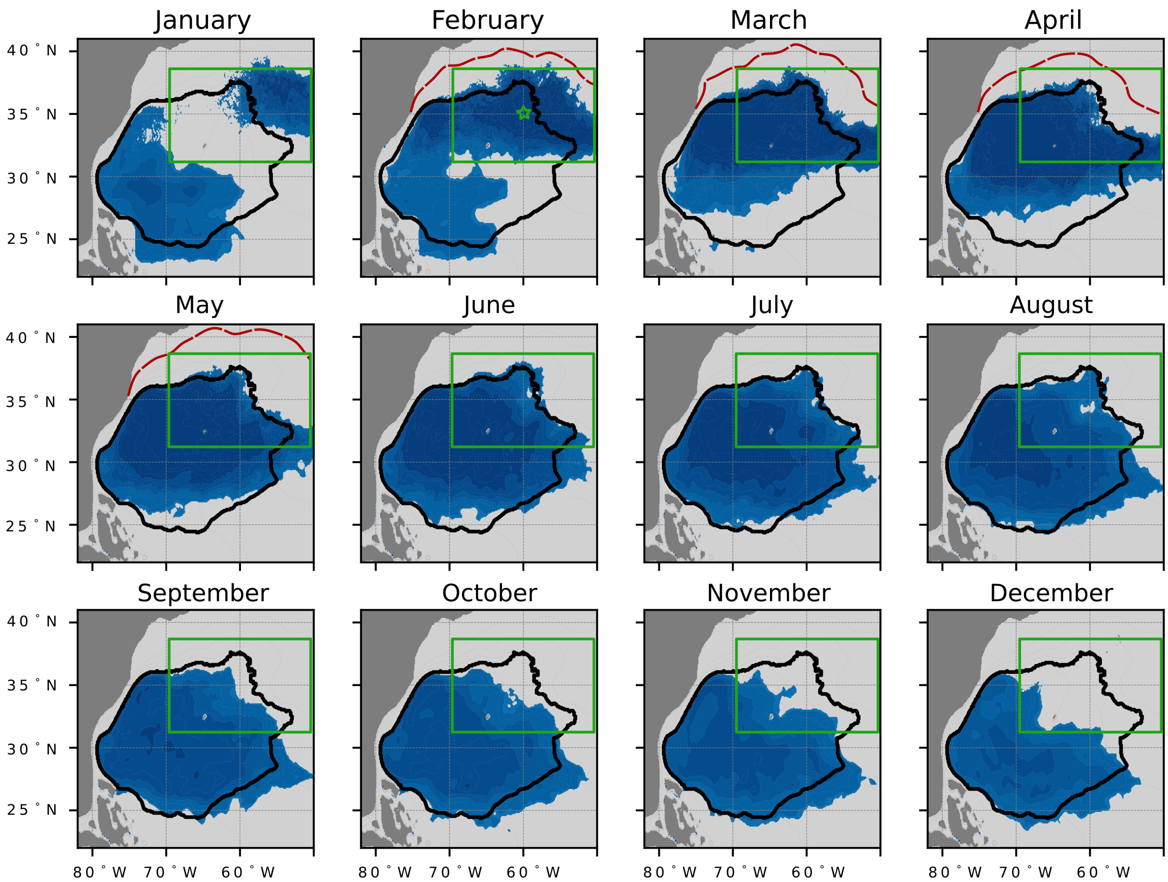


Fig. 5. Monthly-mean PV of STMW ($\overline{g^{-1}\Pi^{\#}} \leq 3 \times 10^{-10}\mathrm{m}^{-1}\mathrm{s}^{-1}$) on surface $\tilde{b} = -0.25\mathrm{ms}^{-2}$ appears and disappears during year 1967. Colors are given by the colorbar in Figure 4b. The lateral boundary C of the control volume Ω is determined by the annual-mean contour on $\overline{g^{-1}\Pi^{\#}} = 3 \times 10^{-10}\mathrm{m}^{-1}\mathrm{s}^{-1}$, and is plotted as the same black contour in each panel. Red dashed lines outline outcrops of this surface from February to May. The 'formation zone,' bounded by the same green box in each panel, marks the region where STMW forms and erodes. Green star located at 35°N, 60°W in February is where Warren (1972) determined the location of STMW formation.

Nonetheless, to ensure that our analysis remains robust with the choice of C based on a constant PV contour, we repeated the calculations using alternative thresholds, i.e., $g^{-1}\Pi^{\#} \leq \{1.8, 2.0, 2.5\} \times 10^{-10}\,\mathrm{m}^{-1}\,\mathrm{s}^{-1}$. In all cases, we obtained qualitatively similar results, suggesting that using a constant PV contour as the lateral boundary is a valid approach, and that the conclusions of the study are not sensitive to the exact choice of PV threshold (not shown).

4. Evolution of PV budget

In this section, we analyse the budget of the PV equation (20) integrated over the control volume Ω, which provides insights into the dynamics of STMW throughout the entire year 1967. The three-dimensional volume integration over Ω is given by

$$\int_\Omega \mathrm{d}V \stackrel{\text{def}}{=} \int_\Omega \langle\sigma\rangle \mathrm{d}\tilde{x}\mathrm{d}\tilde{y}\mathrm{d}\tilde{b}\,. \tag{22}$$

We apply this integration to (20) to study the dynamics of PV. The temporal evolution of the control volume and the integrated PV are shown in Figure 6a and Figure 6b, respectively. Figure 6c shows the time series of the integrated terms from (20). The PV tendency $\Pi^{\#}_{\tilde{t}}$ in Figure 6d is reproduced by summing up the other terms in (20).

A comparison between Figures 6a and 6b shows that a decrease in PV within the STMW is associated with an increase in its volume, and vice versa. As the volume increases from February to early May, PV decreases to its annual minimum. Subsequently, as the volume gradually declines, PV slowly increases. These results confirm that the selected control volume in the core layer effectively captures the key characteristics of STMW formation and erosion.

However, these results do not imply that the total PV is conserved in the control volume, as assumed in classical Geophysical Fluid Dynamics (GFD) analysis. The volume-integrated $\Pi^{\#}$ in Figure 6b reduces to an integral of $f + \hat{v}_{\tilde{x}} - \hat{u}_{\tilde{y}}$ over a fixed surface $\mathrm{d}\tilde{x}\mathrm{d}\tilde{y}$ and a constant differential in $\tilde{b}$. Thus, the time series in Figure 6b represents variations of the relative vorticity content $[\int_\Omega (\hat{v}_{\tilde{x}} - \hat{u}_{\tilde{y}})\mathrm{d}\tilde{x}\mathrm{d}\tilde{y}\mathrm{d}\tilde{b}]_{\tilde{t}}$, offset by a constant given by the planetary vorticity. While the STMW volume decreases during the erosion phase, the relative vorticity content increases. This behavior contradicts the classical GFD expectation that relative vorticity should also decrease in order to conserve PV (Vallis 2017). It is possible that topographic effects, such as the New England seamount chain (Chassignet et al. 2023), may generate secondary circulations that advect relative vorticity into the control volume, counteracting the expected decrease from the classical vortex-tube stretching argument. We leave a detailed examination on this for consideration elsewhere.

In the formulation (20), the planetary vorticity component contributes to the volume-integrated $-\Pi^{\#}\boldsymbol{\nabla}\cdot\hat{\mathbf{u}}$ and $\boldsymbol{\nabla}\cdot\mathbf{J}^{\#}$, and dominates over the relative vorticity contribution (Figure A1). This planetary vorticity effect is associated with temporal variations in layer thickness (or volume) across latitudes within the control volume, whereas the contribution from the relative vorticity becomes significant only during the formation phase when the local circulation is strengthened.

Combining these two terms gives

$$\int_{\Omega} \left[\boldsymbol{\nabla} \cdot \mathbf{J}^{\#} - \Pi^{\#} \boldsymbol{\nabla} \cdot \hat{\mathbf{u}}\right] \mathrm{d}V = \int_{\Omega} \hat{\mathbf{u}} \cdot \left(\Pi^{\#}_{\tilde{x}}, \Pi^{\#}_{\tilde{y}}, \Pi^{\#}_{\tilde{b}}\right) \mathrm{d}V\,. \tag{23}$$

The right-hand side of the equation, plotted in gray in Figure 6c, can be interpreted as the net thickness transport into Ω along the direction of the non-thickness-weighted gradient of PV. Its temporal variation is consistent with that of the total thickness shown in Figure 6a, increasing from January to early May and gradually decreasing thereafter. The positive values indicate that $\hat{\mathbf{u}}$ and $(\Pi^{\#}_{\tilde{x}}, \Pi^{\#}_{\tilde{y}}, \Pi^{\#}_{\tilde{b}})$ are of the same signs, meaning that $\hat{\mathbf{u}}$ transports the volume up the PV gradient, leading to an increase in thickness inside the control volume. In other words, as the PV is transported outward by $\hat{\mathbf{u}}$, the volume is increased inwards.

After June, the diabatic velocity $\hat{\varpi}$ is negligible as the outcropping migrates northwards. During this period, the TWA velocity divergence is persistently positive ($\boldsymbol{\nabla} \cdot \hat{\mathbf{u}} > 0$). This can be indicated by the difference between the gray and orange lines, representing the term $\int_{\Omega} \Pi^{\#} \boldsymbol{\nabla} \cdot \hat{\mathbf{u}}\, \mathrm{d}V = \int_{\Omega} \boldsymbol{\nabla} \cdot \mathbf{J}^{\#}\, \mathrm{d}V - \int_{\Omega} \hat{\mathbf{u}} \cdot \left(\Pi^{\#}_{\tilde{x}}, \Pi^{\#}_{\tilde{y}}, \Pi^{\#}_{\tilde{b}}\right) \mathrm{d}V$. Consequently, $\langle \sigma \rangle^{-1} \langle \sigma \rangle_{\tilde{t}} = -\boldsymbol{\nabla} \cdot \hat{\mathbf{u}} < 0$, and the layer thickness continuously reduces in time. This is consistent with the decreasing total thickness from June to November illustrated in Figures 3 and 6a.

Next, we investigate the role of the TWA velocity $\hat{\mathbf{u}}$ on the evolution of the integrated PV. As shown in Figure 6c, the integrated divergence of the residual-mean PV flux (orange line), $\int_{\Omega} \boldsymbol{\nabla} \cdot \mathbf{J}^{\#} \mathrm{d}V$, remains positive throughout the entire year. Based on the divergence theorem that the integrated divergence of a flux over a volume is equivalent to the net flux out of this volume, $\hat{\mathbf{u}}$ generally transports PV outward, leading to a PV extraction. This transport is globally up the PV gradient, because the PV minimum is located within the control volume from an annual-mean perspective. However, the direction does not always hold in time or space, which will be further discussed in the following section.

The residual-eddy PV flux $\mathbf{F}^{\#}$ is along the lateral direction and does not cross $\tilde{b}$-surfaces. The divergences of the E-P fluxes in $\mathbf{F}^{\#}$ accelerate/decelerate the lateral residual-mean flow $(\hat{u}, \hat{v})$ via the lateral eddy momentum flux and the vertical eddy form stress (Stanley et al. 2020; Maddison and Marshall 2013; Poulsen et al. 2019; Uchida et al. 2022). This eddy effect is reflected onto the

PV transport and is quantified by the divergence $\nabla \cdot \mathbf{F}^{\#}$, represented by the green line in Figure 6c. The negative values of $\int_\Omega \nabla \cdot \mathbf{F}^{\#} \mathrm{d}V$ indicate that eddies tend to transport PV into the control volume in an integrated sense. High PVs are transported down the gradient to mix with low PVs, leading to an increase of PV inside the control volume.

The diabatic flux $\mathbf{G}^{\#}$ and its divergence are at least an order of magnitude smaller than those of $\mathbf{J}^{\#}$ and $\mathbf{F}^{\#}$. The positive divergence (cyan line) indicates that $\mathbf{G}^{\#}$ transports PV out of the control volume, and is only noticeable during ventilation periods between February and May. The diabatic impact on PV transport weakens significantly thereafter as the core layer in STMW is isolated from the surface and subducted into the main thermocline, and contributes to an increase of PV in the control volume from June to December, indicating a persistent PV diffusion through the cap of STMW (Billheimer and Talley 2016).

Unlike $\mathbf{G}^{\#}$, neither $\mathbf{J}^{\#}$ nor $\mathbf{F}^{\#}$ shows obvious seasonality in their time series. Although there are some temporal variations during the formation phase, they are not significant enough to characterize the seasonal cycle of the STMW. Details of how much these fluxes cross the boundary over time are shown in Appendix C. We will discuss the seasonality of the STMW in Section 6.

The total rate of PV change in Ω, shown in light blue in Figure 6d, is reconstructed from the sum of the divergences of the TWA velocity and PV fluxes. Its gradient is negative between January and May, then changes sign with reduced magnitude thereafter, consistent with the evolution of the integrated $\Pi^{\#}$ in Figure 6b.

Overall, our results suggest that the evolution of the PV transport within the control volume is dominated by the residual-mean PV flux, acting up the PV gradient in an integrated sense, and is compensated by the down-gradient residual-eddy PV flux. This finding is consistent with the results by Deremble and Dewar (2013). The effect of the diabatic PV flux is generally weaker than that of the other two PV fluxes, but contributes to a decreasing PV during the formation period and slightly increasing PV during the erosion period.

5. Regional mechanisms

In this section, we focus on the mean-eddy interactions and investigate the dynamical mechanisms of STMW maintenance in different regions in the North Atlantic. We approach it by taking an

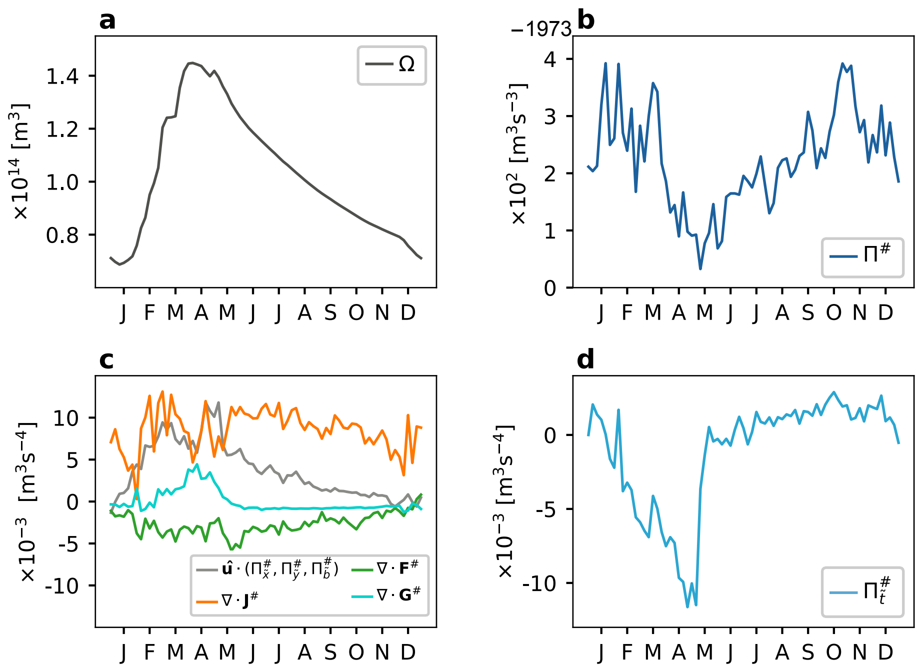


FIG. 6. Time series of integrated variables over the control volume Ω in the core layer. (a) Total volume $\int_\Omega \mathrm{d}V$; (b) Volume integrated PV $\int_\Omega \Pi^\# \mathrm{d}V$; (c) Volume integrated terms of (20): Velocity weighted by the PV-gradient [gray; Eq. (23)], divergences of the residual-mean PV flux (orange), the residual-eddy PV flux (green), and the diabatic PV flux (cyan); (d) Estimated tendency term reproduced by using the PV-gradient weighted velocity, the residual-eddy PV flux divergence and the diabatic PV flux divergence.

annual time-averaging of the integrated PV equation (20) and examining the contribution of PV fluxes to the control volume from a spatial perspective.

By assuming an adiabatic layer ($\hat{\varpi} \sim 0$) in the annually averaged sense, we obtain

$$\int_\Omega \overline{\Pi^\#_{\tilde{t}}} \mathrm{d}V = -\int_\Omega \overline{-\Pi^\# \boldsymbol{\nabla}_h \cdot \hat{\mathbf{u}}_h + \boldsymbol{\nabla}_h \cdot \mathbf{J}^\#_h + \boldsymbol{\nabla}_h \cdot \mathbf{F}^\# + \boldsymbol{\nabla}_h \cdot \mathbf{G}^\#_h} \mathrm{d}V\,, \tag{24}$$

where the subscript h denotes the horizontal (lateral) components of the divergence operator and the fluxes. Residual-eddy PV flux $\mathbf{F}^\#$ is lateral by construction so does not have the subscript. Based on the divergence theorem, the volume integral on the RHS is equivalent to the sum of the fluxes across the lateral boundary surface $\mathcal{S}$ of the control volume Ω.

$$\int_\Omega \overline{-\Pi^\# \boldsymbol{\nabla}_h \cdot \hat{\mathbf{u}}_h + \boldsymbol{\nabla}_h \cdot \mathbf{J}^\#_h + \boldsymbol{\nabla}_h \cdot \mathbf{F}^\# + \boldsymbol{\nabla}_h \cdot \mathbf{G}^\#_h} \mathrm{d}V = \int_\mathcal{S} \left[\overline{-\Pi^\# \hat{\mathbf{u}}_h} + \overline{\mathbf{J}^\#_h} + \overline{\mathbf{F}^\#} + \overline{\mathbf{G}^\#_h}\right] \cdot \mathbf{n} \mathrm{d}\mathcal{S}\,, \tag{25}$$

where $\mathbf{n}$ is the outward-point normal unit vector on surface $\mathcal{S}$. Here, $\overline{\Pi^{\#}}$ is assumed to be nearly homogeneous within Ω, so that the volume integral of the PV-weighted divergence term $\overline{\Pi^{\#}\boldsymbol{\nabla}_h \cdot \hat{\mathbf{u}}_h}$ can be approximated as $\boldsymbol{\nabla}_h \cdot (\overline{\Pi^{\#}}\overline{\hat{\mathbf{u}}_h})$ within Ω (light turquoise line). The lateral surface integral of the time-mean control volume is given by $\int_{\mathcal{S}} \mathrm{d}\mathcal{S} = \oint_C \overline{\langle\sigma\rangle}\mathrm{d}l$, leading to

$$\int_{\mathcal{S}} \left[\overline{-\Pi^{\#}\hat{\mathbf{u}}_h} + \overline{\mathbf{J}_h^{\#}} + \overline{\mathbf{F}^{\#}} + \overline{\mathbf{G}_h^{\#}}\right] \cdot \mathbf{n}\mathrm{d}\mathcal{S} = \oint_C \left[\overline{-\langle\sigma\rangle\Pi^{\#}\hat{\mathbf{u}}_h} + \overline{\langle\sigma\rangle\mathbf{J}_h^{\#}} + \overline{\langle\sigma\rangle\mathbf{F}^{\#}} + \overline{\langle\sigma\rangle\mathbf{G}_h^{\#}}\right] \cdot \mathbf{n}\mathrm{d}l\,, \tag{26}$$

where $\overline{\langle\sigma\rangle}$ is the annual-mean ensemble-mean layer thickness. Thus, the LHS of (24) can be expressed as a line integral of the sum of thickness-weighted fluxes along the closed (lateral) contour C (see black contours in Figure 4),

$$\int_{\Omega} \overline{\Pi_{\tilde{t}}^{\#}}\mathrm{d}V = -\oint_C \left[\overline{-\langle\sigma\rangle\Pi^{\#}\hat{\mathbf{u}}_h} + \overline{\langle\sigma\rangle\mathbf{J}_h^{\#}} + \overline{\langle\sigma\rangle\mathbf{F}^{\#}} + \overline{\langle\sigma\rangle\mathbf{G}_h^{\#}}\right] \cdot \mathbf{n}\mathrm{d}l\,. \tag{27}$$

The thickness-weighted fluxes normal to the contour C are cumulatively integrated in a counter-clockwise direction in Figures 7 and 8. If, within a segment, the line gradient of a flux is positive, then the flux transports PV out of the control volume. Since the annual mean PV is lower inside the control volume, an outward transport is directed up the PV gradient and indicates a PV extraction from STMW, and vice versa. The residual plotted in a dotted line is the RHS of (27), and represents an approximation of $\overline{\Pi_{\tilde{t}}^{\#}}$. Its line integral should ideally be continuous at p_1 since $\Pi^{\#}$ is expected to be homogeneous within the control volume from a time-mean perspective. We attribute the observed discrepancy to unavoidable discretization errors, which will be discussed in detail in Section 7.

The contribution of the diabatic PV flux $\mathbf{G}^{\#}$ (cyan line) is negligible since the layer is nearly adiabatic under an annual averaging. Thus, we focus on $\mathbf{F}^{\#}$ (green line) and $\mathbf{J}^{\#}$ (orange line) and discuss their roles on PV transport in different regions. We label critical locations on the contour C by p-indices shown in Figure 4b, and split C into three segments as in the following subsections.

a. Along-coast/Upstream of the Gulf Stream

The most positive gradient is given by the residual-mean PV flux $\mathbf{J}^{\#}$ between locations p_6 and p_3. This segment of the contour follows the Florida current along the coast between 30°N and 35°N (p_6 - p_5) and then upstream of GS after passing the separation point (p_5 - p_3). The result

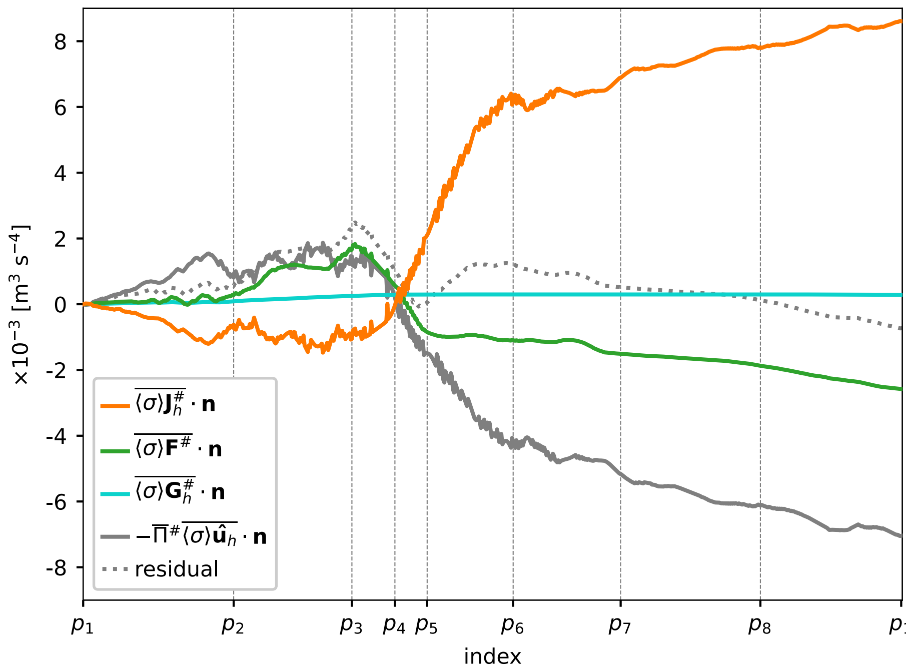


FIG. 7. Thickness-weighted PV fluxes are accumulated along the contour C from p_1 (see Figure 4b) in a counter-clockwise direction. Line colours used to represent the PV fluxes are consistent with those used in Figure 6. The unit vector $\mathbf{n}$ along C points out of the control volume. Indices p_n's are in correspondence with those labelled in Figure 4b, and are locations where the plotted terms change signs or strength. The PV-weighted TWA divergence is approximated by the contour integration of $\overline{\Pi^{\#}}\overline{\langle\sigma\rangle\hat{\mathbf{u}}_h}$, plotted in grey. The residual term (dotted line) is the sum of other lines, and is the estimation of the annual mean PV tendency, $\overline{\Pi^{\#}_{\tilde{t}}}$ which is expected to approach zero.

suggests that the strongest PV extraction is due to $\mathbf{J}^{\#}$ and occurs in the boundary current region. The underlying mechanism is that in this region the strong cross-jet flow intensifies the jet and sharpens the PV gradients. The result here also suggests that the strength of the Florida current as well as the upstream GS may be a key factor in determining the location of STMW, potentially offering an explanation for its persistence throughout the year. Moreover, it may provide insight into why mode water masses consistently form near western boundary currents (Hanawa and Talley 2001; Tsubouchi et al. 2016).

Along the same segment, the negative gradient of the accumulated residual-eddy PV flux $\mathbf{F}^{\#}$ suggests that $\mathbf{F}^{\#}$ transports PV down the gradient into the control volume. $\mathbf{F}^{\#}$ is relatively weak between p_6 and p_5 as eddies plays a secondary role along the coast. It becomes more prominent between p_5 and p_3 since eddy activity is enhanced after the GS detaches from the coast near Cape Hatteras.

b. Northern Sargasso Sea

The segment of contour between locations p_3 and p_1 lies in the northern Sargasso Sea and is south of the GS, where STMW forms and erodes (Maze et al. 2013; Billheimer and Talley 2016). From an annual-mean perspective, the residual-eddy flux $\mathbf{F}^{\#}$ contributes to a decrease of PV within the control volume, indicating a formation of STMW, and the residual-mean flux $\mathbf{J}^{\#}$ contributes to an increase of PV, indicating an erosion of STMW.

Along this segment, both the residual-mean and residual-eddy PV fluxes reverse direction in terms of transporting the PV compared to those in the upstream (Section 5*a*). This is because mean-eddy interactions show distinct mechanisms in the along-coast region, near the separation point, and in the meandering region (Waterman and Hoskins 2013; Kang and Curchitser 2015). Therefore, our results suggest that STMW's annual cycle is highly correlated with the path and strength of the GS, consistent with previous findings (Stevens et al. 2020; Joyce et al. 2000; Tsubouchi et al. 2016).

It should be noted that, while retaining their respective roles in STMW's formation and erosion, the mean and eddy PV fluxes briefly show seasonal variations along this segment. The two PV fluxes locally reverse their directions w.r.t the boundary normal during the ventilation period as shown in Figure 8. When the low PV water is formed to the north and subsequently migrated southward into the control volume (see January and June in Figure 5) by the flow $\hat{\mathbf{u}}$, a reversed PV gradient is produced along some portions of the segment.

Despite these local inversions, the eddy PV flux continues to counteract the mean PV flux, primarily through the vertical form stress. Figure 8 illustrates the respective contributions of the Reynolds stress, vertical form stress, and partial eddy potential energy (EPE) to the eddy PV flux divergence. (Enthusiastic readers are referred to the E–P flux tensors in Equation 10.) Among these components, the vertical form stress is dominant. During the ventilation phase, it mixes the incoming low PV water with its surroundings, thereby driving the seasonal variability of the eddy PV flux in this region. Its contribution is partially offset by the Reynolds stress, which predominantly transports PV outward throughout the year. The term $\frac{\overline{\zeta'^2}}{2\overline{\sigma}}$ in the E-P flux tensors consists only part of the total EPE (see discussion in Uchida et al. 2022, their Eqs. (A15)–(A17)). Its contribution is negligible compared with those of the Reynolds stress and vertical form stress, suggesting that fluctuations in the buoyancy layer show the weakest influence on the eddy effects by $\mathbf{F}^{\#}$.

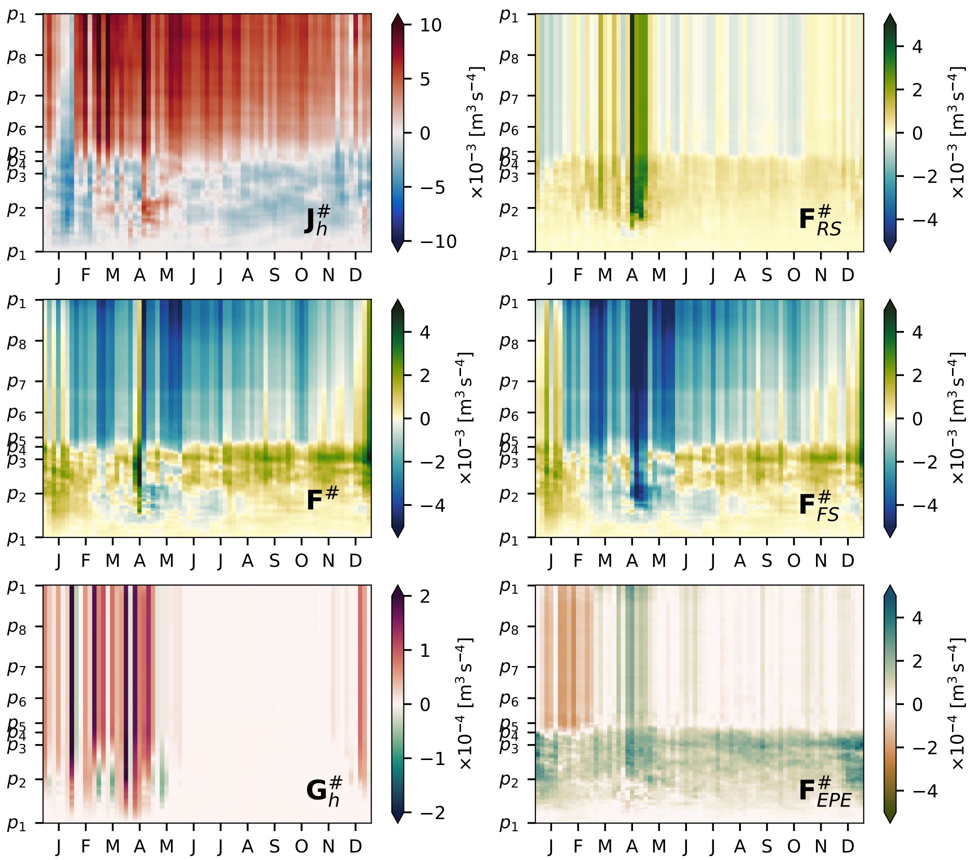


Fig. 8. Thickness-weighted fluxes cumulated from location p_1 along the control volume boundary over the year of 1967. These fluxes are first integrated by the layer thickness in meters, then accumulated from location p_1 on the boundary in a counter-clockwise direction and returning to p_1 (from bottom to top along the y-axis). Their cumulated sums in time are equal to the time series of these PV fluxes in Figure 6c, up to a vertical transport. The residual-eddy PV flux, $\mathbf{F}^{\#}$, is decomposed into its Reynolds stress (RS), vertical form stress (FS) and partial eddy-potential energy (EPE) components, respectively, in the right column.

c. Southern Sargasso Sea

Indices p_1 to p_6 in a clockwise direction along the contour trace the recirculation path of the subtropical gyre. Along this segment of the contour, the extraction of PV by $\mathbf{J}^{\#}$ progressively weakens. A similar finding is reported in Billheimer and Talley (2016), who showed that the destruction rate of STMW is weaker in the southern Sargasso Sea that in the GS region. Although the residual-eddy PV flux $\mathbf{F}^{\#}$ is weaker in magnitude than its mean counterpart $\mathbf{J}^{\#}$ (their spatial fields not shown), its contribution to PV transport is comparable. This is evident when comparing the slopes of the pink and orange lines between locations p_6 and p_1. Because $\mathbf{F}^{\#}$ tends to align more closely with $\mathbf{n}$, the residual eddies transport PV down the gradient and erode STMW more effectively.

6. Role of bolus transport on STMW formation

In this section, we discuss the eddy PV transport by the bolus eddy velocity and its relation to STMW formation to examine its seasonality. The lateral TWA velocity $\hat{\mathbf{u}}_h$ and the lateral ensemble-mean velocity $\langle \mathbf{u} \rangle_h$ are related via a bolus velocity $\mathbf{u}^*$ defined as (17), then the lateral residual-mean PV flux can be decomposed into a lateral ensemble-mean PV flux and a bolus PV flux, respectively,

$$\mathbf{J}_h^{\#} = \langle \mathbf{u} \rangle_h \Pi^{\#} + \mathbf{u}^* \Pi^{\#} . \tag{28}$$

The bolus PV flux $\mathbf{u}^*\Pi^{\#}$ quantifies the lateral eddy advection of PV due to thickness variations.

Figure 9 illustrates the monthly fields of the magnitudes of the bolus PV flux. Starting in February, $\mathbf{u}^*\Pi^{\#}$ begins to emerge in the Northern Sargasso Sea, with the strongest signals observed between 30°N and 35°N (outlined by white dashed lines). This region corresponds to where observation-based studies have detected STMW formation during winter (Maze et al. 2009; Billheimer and Talley 2016), and overlaps with the "formation zone" of the simulated STMW identified in our study (indicated by green boxes in Figure 5).

From July onward, the magnitude of the bolus PV flux declines sharply within the control volume, as the layer thickness is much more uniform. To further investigate, we compare the strength and direction of the ensemble-mean PV flux and the bolus eddy PV flux during two periods—February to June and July to November—in Figure 10. We find that the bolus eddy flux is at least an order of magnitude weaker than the ensemble-mean PV flux and generally transports PV in the opposite direction within the control volume.

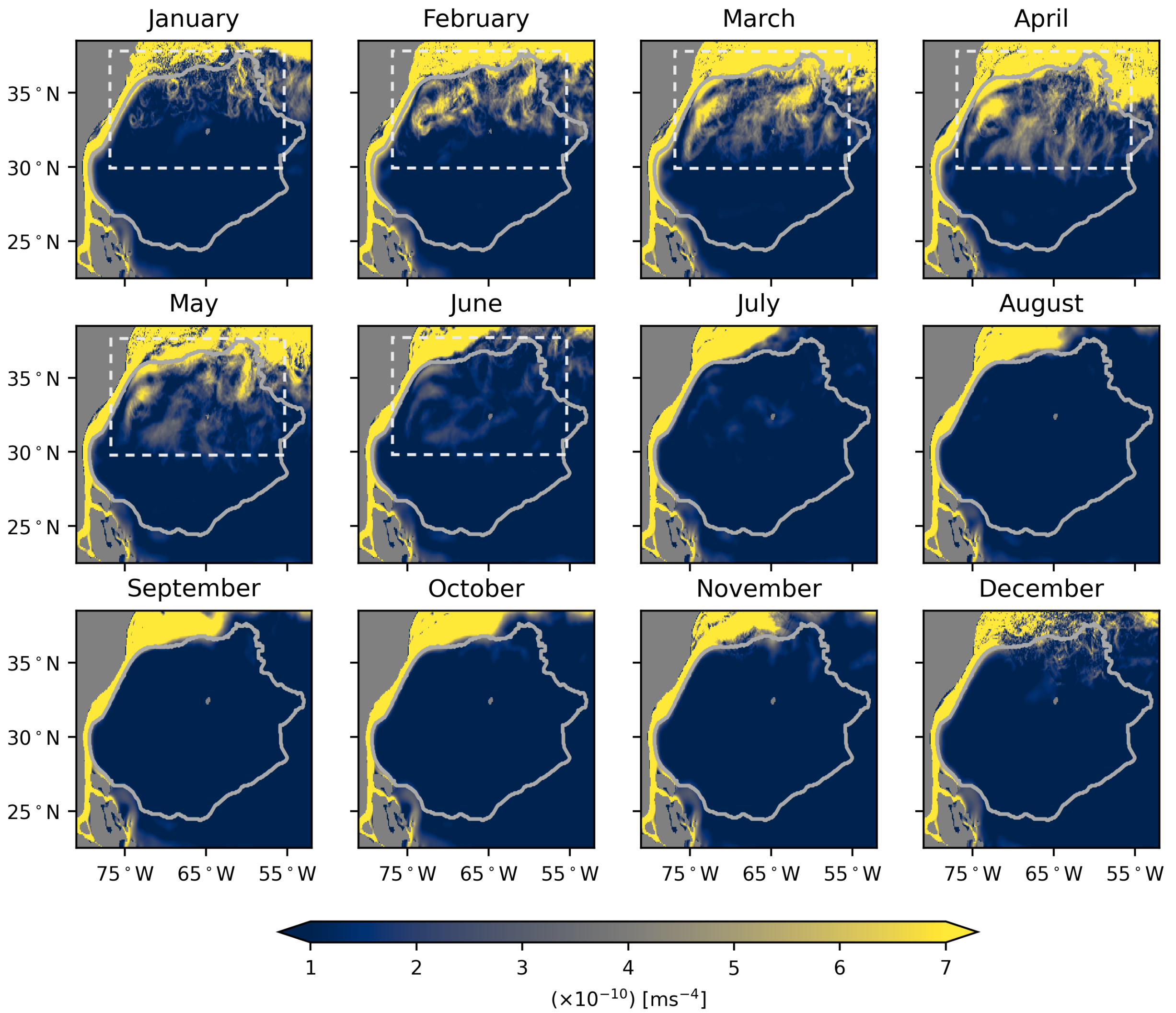


Fig. 9. Magnitudes of the monthly-mean bolus PV flux $\overline{||\mathbf{u}^{*}\Pi^{\#}||}$ on surface $\tilde{b} = -0.25\,\mathrm{m\,s^{-2}}$ in the core layer. Gray contour in each panel is the lateral boundary C of the control volume Ω. Between February and June, the region where the core layer is thickened due to ventilation is outlined by the same white dashed frame. The bolus PV flux inside this region is enhanced significantly in magnitudes due to the intensified baroclinic instability.

Between February and June, the thickness of the core layer within the "formation zone" increases significantly, as indicated by the vertical displacement of the red slope between 30°N and 35°N in Figure 3, leading to enhanced baroclinic instability. As a consequence, an intensified northeastward bolus eddy transport of thickness and PV is observed between 28°N and 36°N in Figure 10, and partially cancels out the southwestward transport associated with $\langle\mathbf{u}_h\rangle$. Between July and November, the bolus PV flux becomes negligible compared to its ensemble-mean counterpart, and the opposing transport directions of the two fluxes are no longer clearly observed.

This time-dependent mean-eddy interaction offers insight into the mechanisms of lateral eddy mixing in STMW formation, which has been identified as crucial in the northern Sargasso Sea near the GS (Billheimer and Talley 2016). Specifically, newly formed low PV water, generated by surface buoyancy loss, is advected southward into the western Sargasso Sea by the large-scale ensemble-mean flow $\langle\mathbf{u}_h\rangle$. It is then actively mixed with pre-existing low PV water via an eddy-induced bolus transport, until the PV field in region, that is, the control volume in our study, is well-mixed (Marshall et al. 1999). An analogy of this can be found for the Deacon cell in the Southern Ocean; the upwelling water in isopycnal layers and northward transport by Ekman pumping is counteracted by an bolus-driven overturning (Karoly et al. 1997; Abernathey et al. 2011; Marshall and Speer 2012).

7. Discussion

Before concluding, we discuss the numerical accuracy of our methodology and the challenges associated with closing the PV budget within the control volume Ω. We use instantaneous snapshot outputs instead of the time-averaged ones because the TWA operator does not commute with temporal averaging. Given that we cannot diagnose the 'true' tendency of PV using instantaneous snapshots every five days, the reproduced PV tendency (Figure 6d) does not exactly align with the offline temporal derivative of PV in Figure 6b. The volume integrated PV budget can be recasted as

$$\left(\int_\Omega \Pi^\# \mathrm{d}V\right)_{\tilde{t}} = \left(\int_\Omega \langle\sigma\rangle \Pi^\# \mathrm{d}\tilde{x}\mathrm{d}\tilde{y}\mathrm{d}\tilde{b}\right)_{\tilde{t}} = -\int_\Omega \nabla\cdot(\mathbf{J}^\# + \mathbf{F}^\# + \mathbf{G}^\#)\mathrm{d}V\,, \tag{29}$$

which, in theory, reads cleaner than (20). However, whether it is included implicitly or explicitly in the integrand on the left-hand side, time rate of change in $\langle\sigma\rangle$ cannot be robustly estimated using snapshot outputs every five days. This stems from ζ_t fluctuating widely in time. As such, we have chosen (20) as our formulation to analyze. We attribute these discrepancies to technical challenges in implementing the TWA framework and determining a time-varying control volume (Stanley and Marshall 2022), and temporal aliasing. We outline some of the technical caveats below.

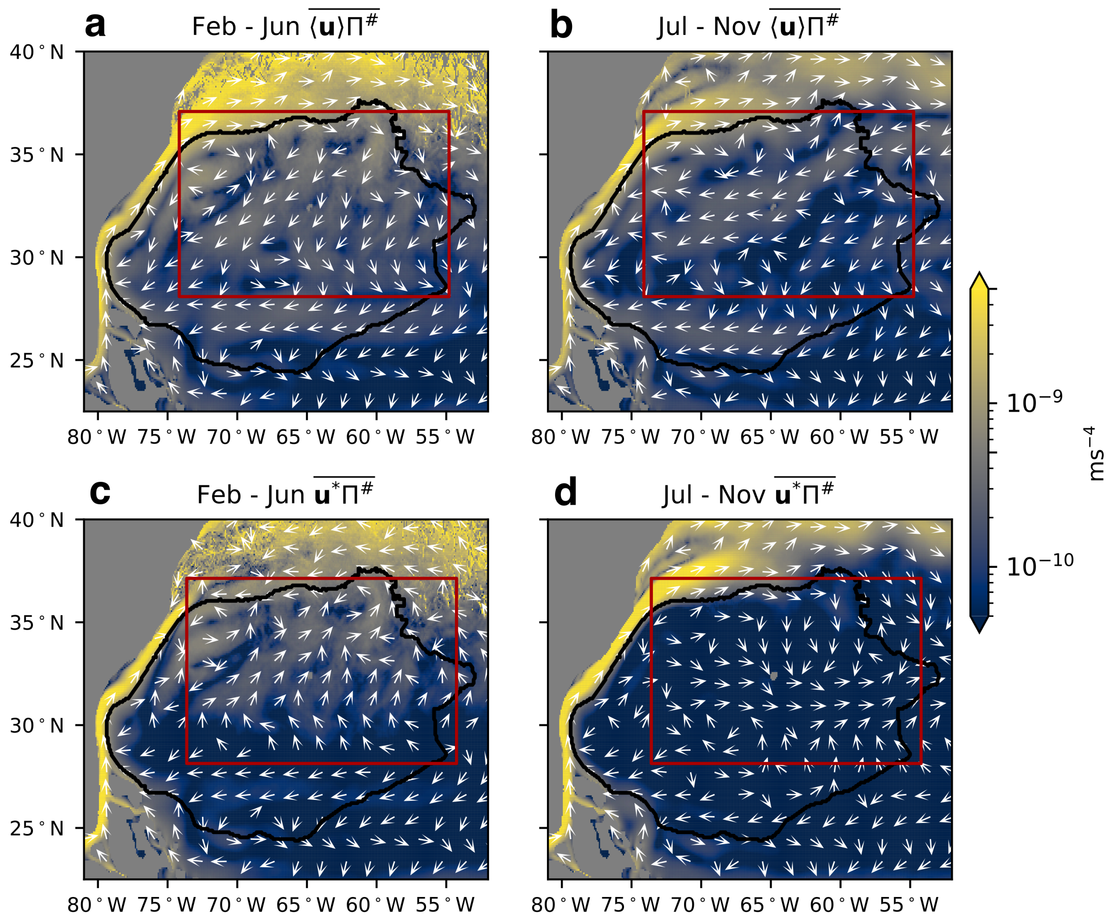


FIG. 10. Time-mean fields of the lateral ensemble-mean PV flux and the bolus PV flux between (a,b) February and June, and (c,d) between July and November. The region where the bolus PV flux intensifies/weakens and is of opposite directions of the ensemble-mean PV flux is outlined by the red frame in each panel.

The residual-mean momentum equations and the thickness equation in (7) cannot be closed to numerical precision in offline diagnostics due to the non-commutability between the TWA operator and space-time gradients, and hence, neither can the PV equation. This is because the layer thickness σ is precisely defined to be spatially and temporally dependent in z-coordinates. When transforming a thickness-weighted variable from z-coordinates to $\tilde{b}$-coordinates, this precision on σ is lost, and discretization errors accumulate in the ensuing computations. For example, even though the divergence operator $\boldsymbol{\nabla}$ is theoretically coordinate-invariant, numerical errors induced by applying the formulation (12) are non-negligible and they further result in a biased integration of $\Pi^{\#}_{\tilde{t}}$. This sensitivity of TWA and thickness equation to offline vertical coordinate transformation has also been reported by Stanley and Marshall (2022).

Furthermore, it is essential to retain the ventilated portion of STMW (Billheimer and Talley 2013, 2016; Maze et al. 2013; Li et al. 2022) to ensure that diabatic effects on STWM formation are captured. However, this introduces two sources of errors in estimating the PV budget within Ω. The accuracy of the methodology will be less promising over the corresponding periods since the buoyancy coordinate becomes increasingly ill-defined in the mixed layer, and $\mathrm{d}\tilde{b}$ and $\mathrm{d}\zeta$ in (2) approach zero towards the surface. We suspect that a better budget closure may be achieved if a layer exhibiting less spatial and temporal variabilities than ours in $\langle\sigma\rangle$ is considered, which does not exist within STWM, and/or if TWA diagnostics are computed online (cf. Ferreira and Marshall 2006; Ringler et al. 2017).

Despite these caveats, we argue that the essential dynamics of PV evolution are well represented by the governing terms in (20), asides from numerical inaccuracies. In Figure 6, $\Pi^{\#}_{\tilde{t}}$ reconstructed by these terms follows the same trend as the integrated $\Pi^{\#}$, and both capture the annual cycle of STMW formation and erosion. As discussed in previous sections, we have demonstrated that the contribution of the governing terms to the PV evolution is consistent with previous theoretical studies (Dewar 1986; Deremble and Dewar 2013).

Lastly, it is possible that the eddy effects are under-represented at our resolution of $1/12°$ (Xu et al. 2025) although this resolution may be sufficient for the purposes of studying the annual cycle of STMW (Prochko and Masumoto 2025). It would be interesting to repeat our analyses using the 48-member $1/50°$ ensemble currently under production (Uchida et al. 2024, their Supplementary Material Figures S2 and S3).

8. Conclusions

In this study, we have investigated the primary mechanisms governing the maintenance of Subtropical Mode Water (STMW) in the North Atlantic. We have focused on the interactions between the mean flow and eddies, and their impacts on the annual life cycle of STMW. We have used an ensemble of 48 partially air-sea coupled North Atlantic simulations at $1/12°$ resolution. This ensemble approach permits a separation of eddies from a full flow without prescribing any temporal or spatial scale. As both eddies and ensemble-mean variables are spatially and

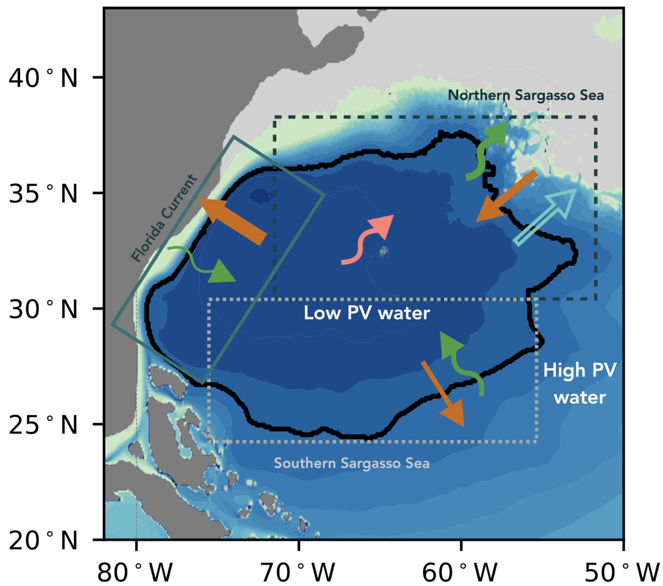


FIG. 11. Schematic illustration of the dynamical mechanisms maintaining the Subtropical Mode Water (STMW) in the North Atlantic. The background shading shows the annual-mean PV $\overline{\Pi^{\#}}$, with darker blue indicating lower PV values. The black contour outlines the low PV pool characterising STMW in its core layer. Distinct mechanisms are found in the three regions outlined with different line styles. Within these regions, orange arrows denote the residual-mean PV flux, and green curved arrows denote the residual-eddy PV flux. Arrow direction indicates the net direction of PV transport, while arrow thickness indicates relative flux magnitude. In the northern Sargasso Sea region, a cyan outlined arrow denotes the outward diabatic PV flux, and a pink curved arrow denotes the bolus PV flux appearing within the bounded region; both are active primarily during the formation phase.

temporally dependent, we were able to explore their full impacts on the dynamics of STMW's annual evolution. By implementing a thickness-weighted averaging (TWA) framework, we have formulated a potential vorticity (PV) equation for $\Pi^{\#}$ in buoyancy layers, analogous to the flux-form Ertel's PV equation. To quantify the dynamics of STMW, we have studied the budget of a low PV pool enclosed by a control volume Ω in the core STMW layer.

Our results highlight that the annual life cycle of STMW is primarily determined by the residual-mean flow, represented by the TWA velocity $\hat{\mathbf{u}}$, whereas the eddies and diabatic forcing play a secondary role. The thickness-weighted ensemble-mean velocity $\hat{\mathbf{u}}$ governs the conservation of the ensemble-mean thickness $\langle\sigma\rangle$ within buoyancy layers. It drives an increase in $\langle\sigma\rangle$ within the control volume between February and May, corresponding to the formation phase of STMW, and a subsequent decrease thereafter, consistent with its erosion phase. On the other hand, the associated residual-mean PV flux $\mathbf{J}^{\#}$ does not show a pronounced seasonal variability. In an integrated and annually-averaged sense, $\mathbf{J}^{\#}$ extracts PV from the control volume through an upgradient advection. The strongest outward PV flux occurs along the coastline, where the PV gradient is the sharpest, and follows the Gulf Stream until the separation point (29°N to 36°N). This persistent extraction of PV throughout the year may provide an explanation for the existence of STMW in the western Sargasso Sea, as well as for the consistent presence of mode waters adjacent to western boundary currents.

These findings also show that the TWA framework can effectively capture the dynamics of STMW, a key climate indicator in the ocean interior, through a time-varying residual mean circulation, even when residual eddies are partially resolved. This underscores a unique strength of the framework: bolus eddy effects are embedded in the residual-mean flow itself, reducing the need to explicitly prognose bolus transports while still accounting for their role in STMW formation.

We show that the bolus eddy transport is strongly linked to the STMW formation phase. The bolus PV flux, $\mathbf{u}^{*}\Pi^{\#}$, intensifies within the control volume between February and June, coinciding with the thickening of the core layer due to surface ventilation. As newly formed low-PV water north of the Gulf Stream is advected into the Sargasso Sea by the ensemble-mean PV flux $\langle\mathbf{u}_h\rangle\Pi^{\#}$, the bolus PV flux associated with the variation of the layer thickness mixes it with pre-existing low-PV water in the control volume by counteracting the ensemble-mean PV flux. This eddy-mean counteraction flattens the tilted buoyancy layers and leads to a nearly homogeneous PV field within the control volume. Since the residual-mean velocity inherently represents the cancellation between the bolus and ensemble-mean velocities, seasonal variations driven by the bolus eddy transport are not directly reflected in the temporal evolution of the residual-mean PV flux.

The eddy forcing in the PV equation is represented by the residual-eddy PV flux $\mathbf{F}^{\#}$, which quantifies the effects of residual eddy momentum fluxes and vertical eddy form stress on PV transport. The divergence of $\mathbf{F}^{\#}$ is approximately one order of magnitude smaller than that of its residual-mean counterpart, $\mathbf{F}^{\#}$. Despite this, $\mathbf{F}^{\#}$ plays a crucial role in the PV evolution within the control volume, and thus in the STMW's life cycle, since it consistently acts to counterbalance the transport by $\mathbf{J}^{\#}$. The integrated residual-eddy PV flux $\mathbf{F}^{\#}$ over the control volume does not have a strong seasonality, similar to the residual-mean PV flux $\mathbf{J}^{\#}$. It tends to transport PV down the gradient, resulting in a net increase of PV within the control volume. This behavior, as well as that by the bolus eddy PV flux, is consistent with the classical assumption of down-gradient eddy transport (Green 1970; Rhines and Young 1982). To rigorously determine whether eddy mixing — whether by residual eddies or bolus eddies — acts down-gradient or up-gradient in a full spatio-temporal sense, it would be necessary to explicitly compute the eddy diffusivity (e.g. Bachman et al. 2020; Haigh et al. 2021; Uchida et al. 2023; Deremble et al. 2023). This question remains important, especially given that many mesoscale eddy parameterizations, such as the Gent-McWilliams' skew diffusion (GM; Gent and Mcwilliams 1990; Griffies 2018) and its extensions (e.g. Treguier et al. 1997; Aiki et al. 2004; Cessi 2008; Marshall et al. 2012; Mak et al. 2023; Torres et al. 2025), are developed around bolus eddies and based on the assumption of down-gradient eddy transport.

We have diagnosed the spatial contributions of the residual-mean and residual-eddy PV fluxes by calculating their net fluxes along different segments of the lateral contour in an annually averaged sense. The contribution of the diabatic PV flux $\mathbf{G}^{\#}$ is found to be small, as the buoyancy layer is nearly adiabatic when averaged over the year; nevertheless, their contribution become noticeable during late winter when the layers outcrop at the surface. Along the coastal segment, eddy effects are secondary because the strong Gulf Stream jet inhibits significant eddy generation. In contrast, the contribution of $\mathbf{F}^{\#}$ becomes most pronounced along the northern boundary of the control volume, situated on the southern flank of the Gulf Stream, where eddy activity intensifies after the Gulf Stream detaches from the coast. This region is also critical for the annual renewal and decay of STMW. The vertical form stress dominates over the Reynolds stress and serves as the primary mechanism counteracting the residual-mean PV flux that transports low PV water during the formation phase. This result further highlights the important role of eddies in both the formation and erosion of STMW. In the southern Sargasso Sea, following the recirculation pathway of the subtropical gyre, the PV transport is approximately balanced between the residual-mean and residual-eddy fluxes. This finding suggests that, even though eddies are relatively weak in this region, their contribution to the maintenance of STMW is nonetheless substantial. Overall, the interaction between the residual-mean flow and residual-eddies, and consequently their impacts on

PV transport, as summarized in Figure 11, varies significantly in space and time, giving rise to a complex and regionally dependent mechanism governing the formation, erosion and maintenance of STMW.

Our work highlights a unique advantage of the ensemble approach, which enables probing into the full spatiotemporal effects of bolus eddies and eddy-mean flow interactions anywhere in the ocean. In contrast to the Southern Ocean, where zonal-mean analyses of the residual-mean flow are commonly employed, the presence of lateral boundaries in the North Atlantic precludes such an approach. Our results, together with previous studies on the Antarctic Circumpolar Current (e.g. Marshall and Radko 2003; Abernathey et al. 2011; Munday et al. 2013; Sinha and Abernathey 2016; Bishop et al. 2016; Constantinou and Hogg 2019, and references therein), suggest that the partial cancellation between the Eulerian-mean and bolus-eddy flows, a phenomenon sometimes referred to as 'eddy saturation', is a ubiquitous feature of oceanic jets.

Acknowledgments. We dedicate this study in memory of Prof. William K. Dewar, with deep respect. L. Sun is especially grateful for his mentorship — his intellectual guidance and unwavering support were instrumental throughout the development of this work. We thank the editor Paola Cessi along with David Marshall and two other anonymous reviewers for their valuable comments. W. Dewar was supported through National Science Foundation (NSF) grants OCE-1829856, OCE-1941963, OCE-2023585, and OCE-2123632, and the French 'Make Our Planet Great Again' (MOPGA) program managed by the Agence Nationale de la Recherche under the Programme d'Investissement d'Avenir, reference ANR-18-MPGA-0002. L. Sun acknowledges the support by the latter two grants. T. Uchida acknowledges the NASA award 80NSSC24K1649 during his time in the U.S. and the Moscow Institute of Physics and Technology Development Program (Priority-2030) upon moving to Russia. A. Poje acknowledges support from the NSF grant OCE-2123633. High-performance computing resources on Cheyenne (`doi:10.5065/D6RX99HX`) used for running the Chaocean ensemble of the North Atlantic were provided by NCAR's Computational and Information Systems Laboratory, sponsored by NSF, under the university large allocation UFSU0011. We are grateful to Qian Li and Yueyang Lu for their insightful discussions. Our study could not have been possible without the production of the NA12 ensemble dataset by Quentin Jamet, to whom we are indebted, and the `xgcm` Python package (Abernathey et al. 2021). We would like to extend our gratitude to Edward Peirce and Kelly Hirai for maintaining the FSU cluster on which the data were analyzed.

Data availability statement. EN.4.2.2 data were obtained from `https://www.metoffice.gov.uk/hadobs/en4/` and are © British Crown Copyright, Met Office, 2013, provided under a Non-Commercial Government Licence (`http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence/version/2/`; Good et al. 2013). The NA12 ensemble outputs are available via the Florida State University (FSU) cluster (`http://ocean.fsu.edu/~qjamet/share/data/twa_Uchida/`).

APPENDIX A

TWA PV conservation equation in buoyancy coordinates

Subtracting the cross derivatives of (7a) and (7b) and using the TWA PV definition (14) leads to

$$\Big(\langle\sigma\rangle\Pi^{\#}\Big)_{\tilde{t}}+\Big(\langle\sigma\rangle\hat{u}\Pi^{\#}+\boldsymbol{\nabla}\boldsymbol{\cdot}\mathbf{E}^{v}-\hat{\mathscr{Y}}+\hat{v}_{\tilde{b}}\hat{\varpi}\Big)_{\tilde{x}}+\Big(\langle\sigma\rangle\hat{v}\Pi^{\#}-\boldsymbol{\nabla}\boldsymbol{\cdot}\mathbf{E}^{u}+\hat{\mathscr{X}}-\hat{u}_{\tilde{b}}\hat{\varpi}\Big)_{\tilde{y}}=0\,. \tag{A1}$$

By applying the chain rule, the tendency term becomes

$$\Big(\langle\sigma\rangle\Pi^{\#}\Big)_{\tilde{t}}=\langle\sigma\rangle_{\tilde{t}}\Pi^{\#}+\langle\sigma\rangle\Pi^{\#}_{\tilde{t}}\,. \tag{A2}$$

Next, we apply (7c) to replace $\langle\sigma\rangle_{\tilde{t}}$ in (A2), and rewrite (A1) as

$$\langle\sigma\rangle\Pi^{\#}_{\tilde{t}}-\langle\sigma\rangle\Pi^{\#}\boldsymbol{\nabla}\boldsymbol{\cdot}\hat{\mathbf{u}}+\Big(\langle\sigma\rangle\hat{u}\Pi^{\#}+\boldsymbol{\nabla}\boldsymbol{\cdot}\mathbf{E}^{v}-\hat{\mathscr{Y}}+\hat{v}_{\tilde{b}}\hat{\varpi}\Big)_{\tilde{x}}+\Big(\langle\sigma\rangle\hat{v}\Pi^{\#}-\boldsymbol{\nabla}\boldsymbol{\cdot}\mathbf{E}^{u}+\hat{\mathscr{X}}-\hat{u}_{\tilde{b}}\hat{\varpi}\Big)_{\tilde{y}}=0\,. \tag{A3}$$

Then we divide (A3) by $\langle\sigma\rangle$ and apply the divergence operator (12). This leads to the PV conservation form (20)

$$\Pi^{\#}_{t}-\Pi^{\#}\boldsymbol{\nabla}\cdot\hat{\mathbf{u}}+\boldsymbol{\nabla}\cdot\mathbf{J}^{\#}+\boldsymbol{\nabla}\cdot\mathbf{F}^{\#}+\boldsymbol{\nabla}\cdot\mathbf{G}^{\#}=0\,, \tag{A4}$$

with PV fluxes substituted into the equation.

APPENDIX B

TWA Bernoulli potential

By introducing the TWA Bernoulli potential

$$\mathscr{B}\stackrel{\text{def}}{=}\frac{1}{2}(\hat{u}^{2}+\hat{v}^{2})+\bar{m}\,, \tag{B1}$$

we rewrite (7a) and (7b) as

$$\hat{u}_{\tilde{t}}+\mathscr{B}_{\tilde{x}}+\hat{\varpi}\hat{u}_{\tilde{b}}-\hat{v}\langle\sigma\rangle\Pi^{\#}+\boldsymbol{\nabla}\cdot\mathbf{E}^{u}-\hat{\mathscr{X}}=0\,, \tag{B2a}$$

$$\hat{v}_{\tilde{t}}+\mathscr{B}_{\tilde{y}}+\hat{\varpi}\hat{v}_{\tilde{b}}+\hat{u}\langle\sigma\rangle\Pi^{\#}+\boldsymbol{\nabla}\cdot\mathbf{E}^{v}-\hat{\mathscr{Y}}=0\,. \tag{B2b}$$

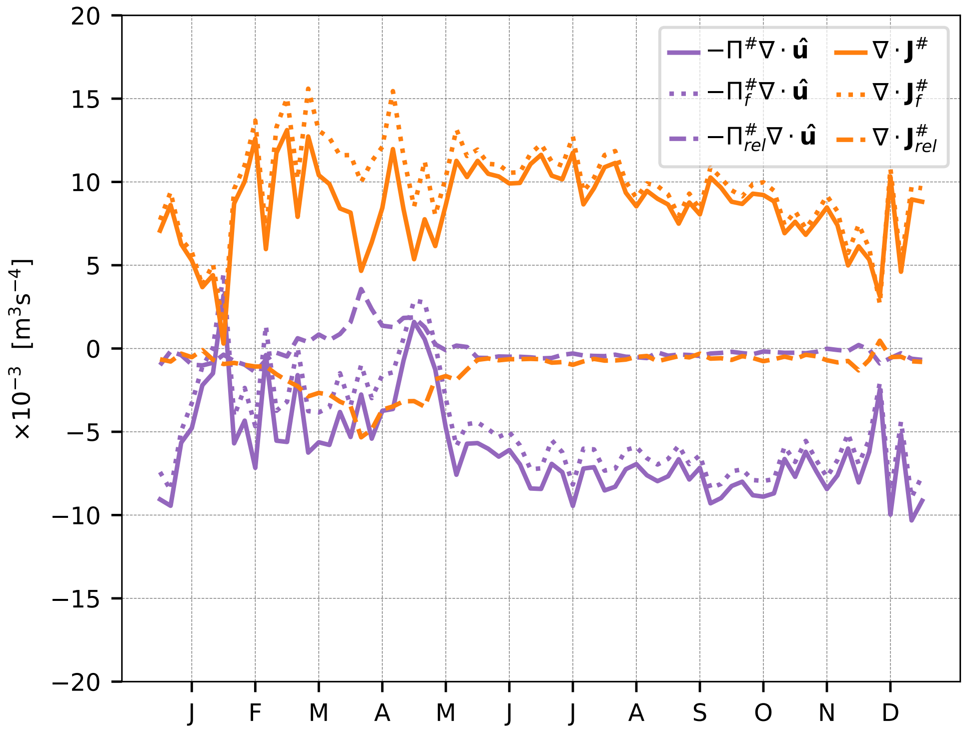


FIG. A1. Time series of the terms $-\Pi^{\#}\boldsymbol{\nabla}\cdot\hat{\mathbf{u}}$ and $\boldsymbol{\nabla}\cdot\mathbf{J}^{\#}$ are shown by the solid purple and orange curves, respectively. The orange curve is identical to that shown in Figure 6c, and their sum is represented by the grey curve in that panel. The dotted curves denote the components associated with the planetary vorticity, whereas the dashed curves denote those associated with the relative vorticity.

We then substitute the PV fluxes back and take a time-averaging,

$$\overline{\hat{u}_{\tilde{t}}} + \overline{\mathcal{B}_{\tilde{x}}} - \overline{\langle\sigma\rangle(J^v + F^v + G^v)} = 0\,, \tag{B3a}$$

$$\overline{\hat{v}_{\tilde{t}}} + \overline{\mathcal{B}_{\tilde{y}}} + \overline{\langle\sigma\rangle(J^u + F^u + G^u)} = 0\,, \tag{B3b}$$

where (J^u, J^v), (F^u, F^v), (Γ^u, Γ^v) are the components of the PV fluxes on $\tilde{b}$-surfaces.

The tendency terms in (B3) become negligible if the time-averaging is taken over a sufficiently long period of time. Then, the time-mean Bernoulli potential can be considered as the streamfunction for a non-divergent flow on a $\tilde{b}$-surface described by the time-mean net PV fluxes,

$$\overline{\mathcal{B}}_{\tilde{x}} \simeq \overline{\langle\sigma\rangle(J^v + F^v + G^v)}\,, \tag{B4a}$$

$$\overline{\mathcal{B}}_{\tilde{y}} \simeq -\overline{\langle\sigma\rangle(J^u + F^u + G^u)}\,. \tag{B4b}$$

One can obtain the time-dependent Bernoulli potential $\mathcal{B}$ directly from (B1). However, this approach is less appropriate in practice, since it requires subtracting large variables of opposite signs to yield a comparatively small residual (Deremble and Dewar 2013; Deremble et al. 2014). A more suitable approach is to compute the time-mean Bernoulli potential $\overline{\mathcal{B}}$ by solving the elliptic equation subject to Neumann boundary conditions,

$$\overline{\mathcal{B}}_{\tilde{x}\tilde{x}} + \overline{\mathcal{B}}_{\tilde{y}\tilde{y}} = \left(\overline{\langle\sigma\rangle(J^v + F^v + G^v)}\right)_{\tilde{x}} - \left(\overline{\langle\sigma\rangle(J^u + F^u + G^u)}\right)_{\tilde{y}}\,. \tag{B5}$$

Thus, the solution remains dynamically consistent with the PV fluxes.

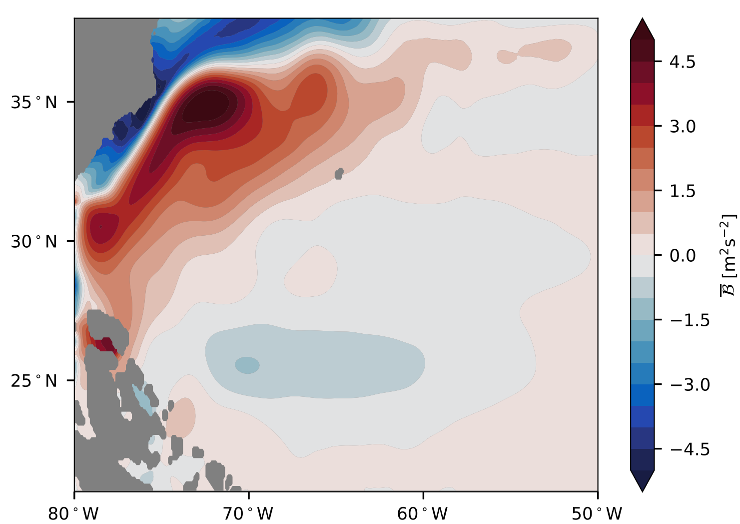
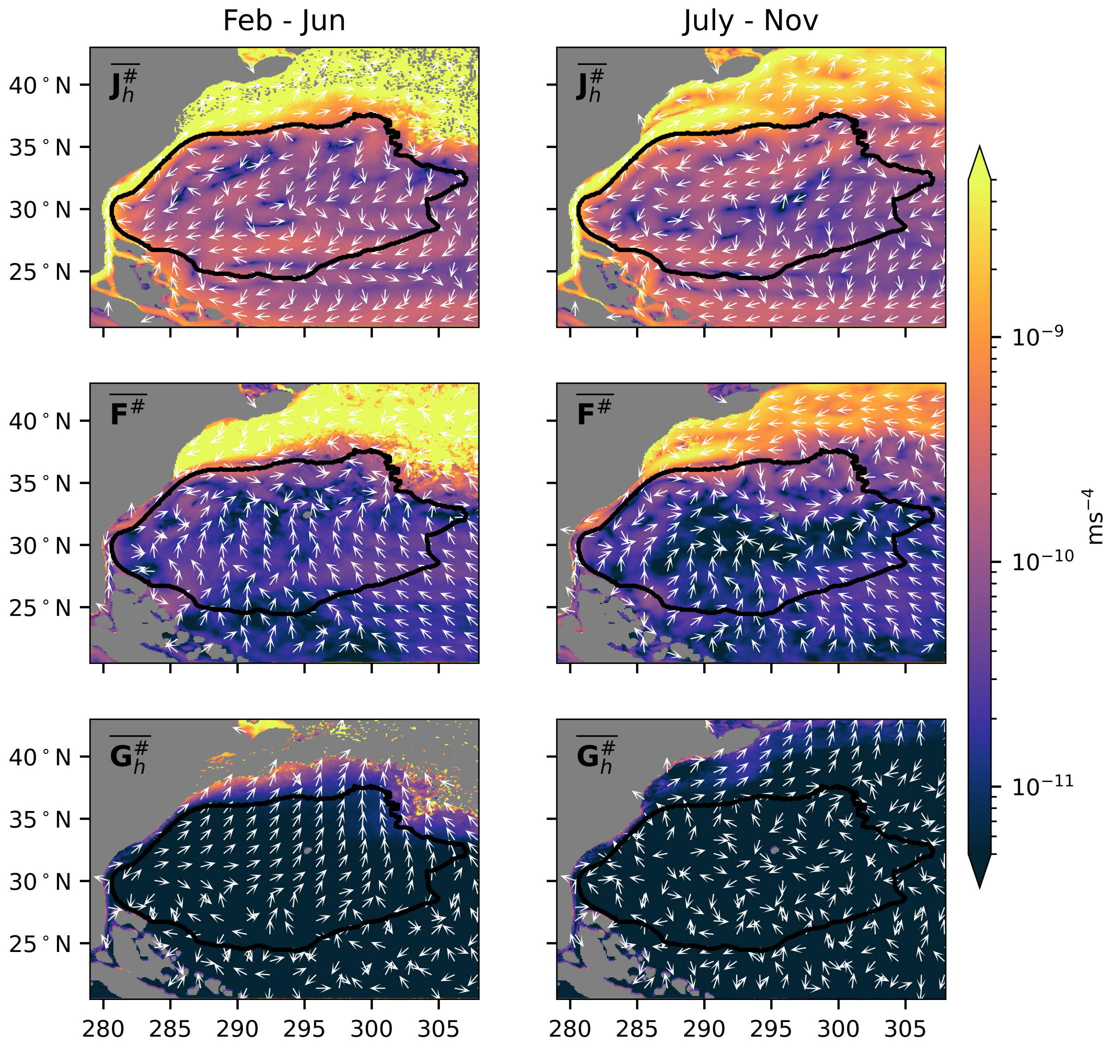


FIG. B1. Time-mean Bernoulli potential field on a $\tilde{b}$-surface.

APPENDIX C

Time evolution of lateral PV fluxes



744 FIG. C1. Spatial fields of the lateral fluxes during the formation and erosion phases of STMW. The colormap
745 indicates the flux magnitude, and the white arrows indicate their direction.

746 References

747 Abernathey, R. P., J. Marshall, and D. Ferreira, 2011: The dependence of Southern Ocean meridional overturning on wind stress. *Journal of Physical Oceanography*, **41 (12)**, 2261–2278, https://doi.org/10.1175/JPO-D-11-023.1.

750 Abernathey, R. P., and Coauthors, 2021: `xgcm`: General Circulation Model Postprocessing with xarray [Software]. URL https://xgcm.readthedocs.io/en/latest/, Zenodo, https://doi.org/10.5281/zenodo.3634752.

Aiki, H., T. Jacobson, and T. Yamagata, 2004: Parameterizing ocean eddy transports from surface to bottom. *Geophysical research letters*, **31 (19)**, https://doi.org/10.1029/2004GL020703.

Aoki, K., 2014: A constraint on the thickness-weighted average equation of motion deduced from energetics. *Journal of Marine Research*, **72**, 355–382, URL https://elischolar.library.yale.edu/cgi/viewcontent.cgi?article=1396&context=journal_of_marine_research.

Bachman, S. D., B. Fox-Kemper, and F. O. Bryan, 2020: A Diagnosis of Anisotropic Eddy Diffusion From a High-Resolution Global Ocean Model. *Journal of Advances in Modeling Earth Systems*, **12 (2)**, https://doi.org/10.1029/2019MS001904.

Billheimer, S., and L. D. Talley, 2013: Near cessation of Eighteen Degree Water renewal in the western North Atlantic in the warm winter of 2011–2012. *Journal of Geophysical Research: Oceans*, **118 (12)**, 6838–6853, https://doi.org/10.1002/2013JC009024.

Billheimer, S., and L. D. Talley, 2016: Annual cycle and destruction of Eighteen Degree Water. *J. Geophys. Res. Oceans*, **121 (9)**, 6604–6617, https://doi.org/10.1002/2016JC011799.

Bishop, S. P., P. R. Gent, F. O. Bryan, A. F. Thompson, M. C. Long, and R. P. Abernathey, 2016: Southern Ocean overturning compensation in an eddy-resolving climate simulation. *Journal of Physical Oceanography*, **46 (5)**, 1575–1592, https://doi.org/10.1175/JPO-D-15-0177.1.

Cessi, P., 2008: An energy-constrained parameterization of eddy buoyancy flux. *Journal of Physical Oceanography*, **38 (8)**, 1807–1819, https://doi.org/10.1175/2007JPO3812.1.

Chassignet, E. P., X. Xu, A. Bozec, and T. Uchida, 2023: Impact of the New England seamount chain on Gulf Stream pathway and variability. *Journal of Physical Oceanography*, **53 (8)**, 1871–1886, https://doi.org/10.1175/JPO-D-23-0008.1.

Cheng, L., J. Zhu, R. Cowley, T. Boyer, and S. Wijffels, 2014: Time, probe type, and temperature variable bias corrections to historical expendable bathythermograph observation. *Journal of Atmospheric and Oceanic Technology*, **31 (08)**, https://doi.org/https://doi.org/10.1175/JTECH-D-13-00197.1.

Constantinou, N. C., and A. M. Hogg, 2019: Eddy saturation of the Southern Ocean: A baroclinic versus barotropic perspective. *Geophysical Research Letters*, **46 (21)**, 12 202–12 212, https://doi.org/10.1029/2019GL084117.

Constantinou, N. C., and A. M. Hogg, 2021: Intrinsic oceanic decadal variability of upper-ocean heat content. *Journal of Climate*, **34 (15)**, 6175–6189, https://doi.org/10.1175/JCLI-D-20-0774.1.

Czaja, A., and U. Hausmann, 2009: Observations of Entry and Exit of Potential Vorticity at the Sea Surface. *Journal of Physical Oceanography*, **39**, https://doi.org/10.1175/2009JPO4024.1.

de Boyer Montégut, C., G. Madec, A. S. Fischer, A. Lazar, and D. Iudicone, 2004: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research: Oceans*, **109 (C12)**, https://doi.org/10.1029/2004JC002378.

Deremble, B., and W. K. Dewar, 2013: Volume and Potential Vorticity Budgets of Eighteen Degree Water. *Journal of Physical Oceanography*, **43 (11)**, 2309–2321, https://doi.org/10.1175/JPO-D-13-052.1.

Deremble, B., T. Uchida, W. K. Dewar, and R. M. Samelson, 2023: Eddy-mean flow interaction with a multiple scale quasi-geostrophic model. *Journal of Advances in Modeling Earth Systems*, **15 (10)**, e2022MS003 572, https://doi.org/10.1029/2022MS003572.

Deremble, B., N. Wienders, and W. K. Dewar, 2013: CheapAML: A simple, atmospheric boundary layer model for use in ocean-only model calculations. *Monthly Weather Review*, **141 (2)**, 809–821, https://doi.org/10.1175/MWR-D-11-00254.1.

Deremble, B., N. Wienders, and W. K. Dewar, 2014: Potential Vorticity Budgets in the North Atlantic Ocean. *Journal of Physical Oceanography*, **44 (1)**, 164–178, https://doi.org/10.1175/JPO-D-13-087.1.

Dewar, W. K., 1986: On the Potential Vorticity Structure of Weakly Ventilated Isopycnals: A Theory of Subtropical Mode Water Maintenance. *Journal of Physical Oceanography*, **16 (7)**, 1204–1216, https://doi.org/10.1175/1520-0485(1986)016⟨1204:OTPVSO⟩2.0.CO;2.

Ferreira, D., and J. Marshall, 2006: Formulation and implementation of a “residual-mean” ocean circulation model. *Ocean Modelling*, **13 (1)**, 86–107, https://doi.org/10.1016/j.ocemod.2005.12.001.

Forget, G., 2010: Mapping ocean observations in a dynamical framework: A 2004–06 ocean atlas. *Journal of Physical Oceanography*, **40 (6)**, 1201–1221, https://doi.org/10.1175/2009JPO4043.1.

Forget, G., G. Maze, M. Buckley, and J. Marshall, 2011: Estimated Seasonal Cycle of North Atlantic Eighteen Degree Water Volume. *Journal of Physical Oceanography*, **41 (2)**, 269–286, https://doi.org/10.1175/2010JPO4257.1.

Gent, P. R., and J. C. Mcwilliams, 1990: Isopycnal Mixing in Ocean Circulation Models. *Journal of Physical Oceanography*, **20 (1)**, 150–155, https://doi.org/10.1175/1520-0485(1990)020⟨0150:IMIOCM⟩2.0.CO;2.

Good, S. A., M. J. Martin, and N. A. Rayner, 2013: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, **118 (12)**, 6704–6716, https://doi.org/10.1002/2013JC009067.

Green, J. S. A., 1970: Transfer properties of the large-scale eddies and the general circulation of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **96 (408)**, 157–185, https://doi.org/10.1002/qj.49709640802.

Griffies, S. M., 2018: *Fundamentals of ocean climate models*. Princeton university press.

Haigh, M., L. Sun, J. C. McWilliams, and P. S. Berloff, 2021: On eddy transport in the ocean. Part I: The diffusion tensor. *Ocean Modelling*, **164**, 101 831, https://doi.org/10.1016/j.ocemod.2021.101831.

Hanawa, K., and L. D. Talley, 2001: Mode waters. *International geophysics*, Vol. 77, Elsevier, 373–386.

Haynes, P. H., and M. E. McIntyre, 1987: On the Evolution of Vorticity and Potential Vorticity in the Presence of Diabatic Heating and Frictional or Other Forces. *Journal of the Atmospheric Sciences*, **44 (5)**, 828–841, https://doi.org/10.1175/1520-0469(1987)044⟨0828:OTEOVA⟩2.0.CO;2.

Jamet, Q., W. K. Dewar, N. Wienders, and B. Deremble, 2019: Spatiotemporal Patterns of Chaos in the Atlantic Overturning Circulation. *Geophysical Research Letters*, **46 (13)**, 7509–7517, https://doi.org/10.1029/2019GL082552.

Jamet, Q., W. K. Dewar, N. Wienders, B. Deremble, S. Close, and T. Penduff, 2020: Locally and remotely forced subtropical AMOC variability: A matter of time scales. *Journal of Climate*, **33 (12)**, 5155–5172, https://doi.org/10.1175/jcli-d-19-0844.1.

Joyce, T. M., 2013: New perspectives on eighteen-degree water formation in the north atlantic. *New Developments in Mode-Water Research: Dynamic and Climatic Effects*, 41–48, https://doi.org/10.1007/978-4-431-54162-2_3.

Joyce, T. M., C. Deser, and M. A. Spall, 2000: The Relation between Decadal Variability of Subtropical Mode Water and the North Atlantic Oscillation. *Journal of Climate*, **13 (14)**, 2550–2569, https://doi.org/10.1175/1520-0442(2000)013⟨2550:TRBDVO⟩2.0.CO;2.

Joyce, T. M., L. N. Thomas, W. K. Dewar, and J. B. Girton, 2013: Eighteen Degree Water formation within the Gulf Stream during CLIMODE. *Deep Sea Research Part II: Topical Studies in Oceanography*, **91**, 1–10, https://doi.org/10.1016/j.dsr2.2013.02.019.

Kang, D., and E. N. Curchitser, 2015: Energetics of eddy–mean flow interactions in the Gulf Stream region. *Journal of Physical Oceanography*, **45 (4)**, 1103–1120, https://doi.org/10.1175/JPO-D-14-0200.1.

Karoly, D. J., P. C. McIntosh, P. Berrisford, T. J. McDougall, and A. C. Hirst, 1997: Similarities of the Deacon cell in the Southern Ocean and Ferrel cells in the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, **123 (538)**, 519–526.

Kwon, Y.-O., and S. C. Riser, 2004: North Atlantic Subtropical Mode Water: A history of ocean-atmosphere interaction 1961–2000. *Geophysical Research Letters*, **31 (19)**, https://doi.org/10.1029/2004GL021116.

Leroux, S., T. Penduff, L. Bessières, J.-M. Molines, J.-M. Brankart, G. Sérazin, and Coauthors, 2018: Intrinsic and atmospherically forced variability of the amoc: Insights from a large-ensemble ocean hindcast. *Journal of Climate*, **31 (3)**, 1183–1203, https://doi.org/10.1175/JCLI-D-17-0163.1.

Li, K., G. Maze, and H. Mercier, 2022: Ekman Transport as the Driver of Extreme Interannual Formation Rates of Eighteen Degree Water. *Journal of Geophysical Research: Oceans*, **127 (1)**, https://doi.org/10.1029/2021JC017696.

Luyten, J. R., J. Pedlosky, and H. Stommel, 1983: The Ventilated Thermocline. *Journal of Physical Oceanography*, **13 (2)**, 292–309, https://doi.org/10.1175/1520-0485(1983)013⟨0292:TVT⟩2.0.CO;2.

Maddison, J. R., and D. P. Marshall, 2013: The Eliassen–Palm flux tensor. *Journal of Fluid Mechanics*, **729**, 69–102, https://doi.org/10.1017/jfm.2013.259.

Mak, J., J. R. Maddison, D. P. Marshall, X. Ruan, Y. Wang, and L. Yeow, 2023: Scale-awareness in an eddy energy constrained mesoscale eddy parameterization. *Journal of Advances in Modeling Earth Systems*, **15 (12)**, e2023MS003 886, https://doi.org/10.1029/2023MS003886.

Marshall, D. P., 2000: Vertical fluxes of potential vorticity and the structure of the thermocline. *Journal of Physical Oceanography*, **30 (12)**, 3102–3112, https://doi.org/10.1175/1520-0485(2000)030⟨3102:VFOPVA⟩2.0.CO;2.

Marshall, D. P., J. R. Maddison, and P. S. Berloff, 2012: A Framework for Parameterizing Eddy Potential Vorticity Fluxes. *Journal of Physical Oceanography*, **42 (4)**, 539–557, https://doi.org/10.1175/JPO-D-11-048.1.

Marshall, D. P., R. G. Williams, and M. M. Lee, 1999: The relation between eddy-induced transport and isopycnic gradients of potential vorticity. *Journal of Physical Oceanography*, **29 (7)**, 1571–1578, https://doi.org/10.1175/1520-0485(1999)029⟨1571:TRBEIT⟩2.0.CO;2.

Marshall, J., C. Hill, L. Perelman, and A. Adcroft, 1997: Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *Journal of Geophysical Research: Oceans*, **102 (C3)**, 5733–5752, https://doi.org/10.1029/96jc02776.

Marshall, J., D. Jamous, and J. Nilsson, 2001: Entry, Flux, and Exit of Potential Vorticity in Ocean Circulation. *J. Phys. Oceanogr.*, **31 (3)**, 777–789, https://doi.org/10.1175/1520-0485(2001)031⟨0777:EFAEOP⟩2.0.CO;2.

Marshall, J., and T. Radko, 2003: Residual-mean solutions for the Antarctic Circumpolar Current and its associated overturning circulation. *Journal of Physical Oceanography*, **33 (11)**, 2341–2354, https://doi.org/10.1175/1520-0485(2003)033⟨2341:RSFTAC⟩2.0.CO;2.

Marshall, J., and K. Speer, 2012: Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience*, **5**, 171–180, https://doi.org/10.1038/ngeo1391.

Maze, G., J. Deshayes, J. Marshall, A.-M. Tréguier, A. Chronis, and L. Vollmer, 2013: Surface vertical PV fluxes and subtropical mode water formation in an eddy-resolving numerical simulation. *Deep Sea Research Part II: Topical Studies in Oceanography*, **91**, 128–138, https://doi.org/10.1016/j.dsr2.2013.02.026.

Maze, G., G. Forget, M. Buckley, J. Marshall, and I. Cerovecki, 2009: Using transformation and formation maps to study the role of air–sea heat fluxes in North Atlantic Eighteen Degree Water formation. *Journal of Physical Oceanography*, **39 (8)**, 1818–1835, https://doi.org/10.1175/2009JPO3985.1.

Maze, G., and J. Marshall, 2011: Diagnosing the Observed Seasonal Cycle of Atlantic Subtropical Mode Water Using Potential Vorticity and Its Attendant Theorems. *Journal of Physical Oceanography*, **41**, 1986–1999, https://doi.org/10.1175/2011JPO4576.1.

Munday, D. R., H. L. Johnson, and D. P. Marshall, 2013: Eddy saturation of equilibrated circumpolar currents. *Journal of Physical Oceanography*, **43 (3)**, 507–532, https://doi.org/10.1175/JPO-D-12-095.1.

Olsina, O., N. Wienders, and W. K. Dewar, 2013: An estimate of the climatology and variability of Eighteen Degree Water potential vorticity forcing. *Deep Sea Research Part II: Topical Studies in Oceanography*, **91**, 84–95, https://doi.org/10.1016/j.dsr2.2013.02.018.

Penduff, T., M. Juza, B. Barnier, J. Zika, W. K. Dewar, A.-M. Treguier, and Coauthors, 2011: Sea level expression of intrinsic and forced ocean variabilities at interannual time scales. *Journal of Climate*, **24 (21)**, 5652–5670, https://doi.org/10.1175/JCLI-D-11-00077.1.

Peng, G., E. P. Chassignet, Y.-O. Kwon, and S. C. Riser, 2006: Investigation of variability of the North Atlantic Subtropical Mode Water using profiling float data and numerical model output. *Ocean Modelling*, **13 (1)**, 65–85, https://doi.org/10.1016/j.ocemod.2005.07.001.

Polton, J. A., and D. P. Marshall, 2003: Understanding the structure of the subtropical thermocline. *Journal of Physical Oceanography*, **33 (6)**, 1240–1249, https://doi.org/10.1175/1520-0485(2003)033⟨1240:UTSOTS⟩2.0.CO;2.

Polton, J. A., and D. P. Marshall, 2007: Overturning cells in the Southern Ocean and subtropical gyres. *Ocean Science*, **3 (1)**, 17–30.

Poulsen, M. B., M. Jochum, J. R. Maddison, D. P. Marshall, and R. Nuterman, 2019: A geometric interpretation of southern ocean eddy form stress. *Journal of Physical Oceanography*, **49 (10)**, 2553–2570, https://doi.org/10.1175/JPO-D-18-0220.1.

Prochko, T., and Y. Masumoto, 2025: Influences of meso-scale and submeso-scale variability on simulated subtropical mode water formation: comparing two different grid-resolutions. *Journal of Oceanography*, 1–20, https://doi.org/10.1007/s10872-025-00774-8.

Rhines, P. B., and W. R. Young, 1982: Homogenization of potential vorticity in planetary gyres. *J. Fluid Mech.*, **122 (-1)**, 347, https://doi.org/10.1017/S0022112082002250.

Ringler, T., J. A. Saenz, P. J. Wolfram, and L. V. Roekel, 2017: A Thickness-Weighted Average Perspective of Force Balance in an Idealized Circumpolar Current. *Journal of Physical Oceanography*, **47 (2)**, 285–302, https://doi.org/10.1175/JPO-D-16-0096.1.

Sinha, A., and R. P. Abernathey, 2016: Time scales of Southern Ocean eddy equilibration. *Journal of Physical Oceanography*, **46 (9)**, 2785–2805, https://doi.org/10.1175/JPO-D-16-0041.1.

Stanley, G. J., T. E. Dowling, M. E. Bradley, and D. P. Marshall, 2020: Ertel Potential Vorticity versus Bernoulli Potential on Approximately Neutral Surfaces in the Antarctic Circumpolar Current. *Journal of Physical Oceanography*, **50 (9)**, 2621–2648, https://doi.org/10.1175/JPO-D-19-0140.1.

Stanley, G. J., and D. P. Marshall, 2022: Why Mean Potential Vorticity Cannot Be Materially Conserved in the Eddying Southern Ocean. *Journal of Physical Oceanography*, **52 (8)**, 1629–1654, https://doi.org/10.1175/JPO-D-21-0195.1.

Stevens, S. W., R. J. Johnson, G. Maze, and N. R. Bates, 2020: A recent decline in North Atlantic subtropical mode water formation. *Nat. Clim. Chang.*, **10 (4)**, 335–341, https://doi.org/10.1038/s41558-020-0722-3.

Torres, R., R. Waldman, G. Madec, C. de Lavergne, R. Séférian, and J. Mak, 2025: An energetically and observationally constrained mesoscale parameterization for ocean climate models. *Journal of Advances in Modeling Earth Systems*, **17 (12)**, e2025MS005 394, https://doi.org/10.1029/2025MS005394.

Treguier, A. M., I. M. Held, and V. D. Larichev, 1997: Parameterization of Quasigeostrophic Eddies in Primitive Equation Ocean Models. *Journal of Physical Oceanography*, **27 (4)**, 567–580, https://doi.org/10.1175/1520-0485(1997)027⟨0567:POQEIP⟩2.0.CO;2.

Tsubouchi, T., T. Suga, and K. Hanawa, 2016: Comparison Study of Subtropical Mode Waters in the World Ocean. *Frontiers in Marine Science*, **3**, https://doi.org/10.3389/fmars.2016.00270.

Uchida, T., D. Balwada, Q. Jamet, W. K. Dewar, B. Deremble, T. Penduff, and J. Le Sommer, 2023: Cautionary tales from the mesoscale eddy transport tensor. *Ocean Modelling*, **182**, 102 172, https://doi.org/10.1016/j.ocemod.2023.102172.

Uchida, T., Q. Jamet, W. K. Dewar, J. Le Sommer, T. Penduff, and D. Balwada, 2022: Diagnosing the Thickness-Weighted Averaged Eddy-Mean Flow Interaction From an Eddying North Atlantic Ensemble: The Eliassen-Palm Flux. *Journal of Advances in Modeling Earth Systems*, **14 (5)**, e2021MS002 866, https://doi.org/10.1029/2021MS002866.

Uchida, T., Q. Jamet, A. C. Poje, N. Wienders, and W. K. Dewar, 2024: Wavelet-based wavenumber spectral estimate of eddy kinetic energy: Application to the North Atlantic. *Ocean Modelling*, **190**, 102 392, https://doi.org/10.1016/j.ocemod.2024.102392.

Uchida, T., Q. Jamet, A. C. Poje, N. Wienders, L. Sun, and W. K. Dewar, 2025: Dynamics and thermodynamics of the Boussinesq North Atlantic eddy kinetic energy spectral budget. *Journal of Advances in Modeling Earth Systems*, **17 (4)**, e2024MS004 781, https://doi.org/10.1029/2024MS004781.

Vallis, G. K., 2017: *Atmospheric and oceanic fluid dynamics*. 2nd ed., Cambridge University Press.

Walin, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the ocean. **34 (2)**, 187, https://doi.org/10.3402/tellusa.v34i2.10801, number: 2 Publisher: Stockholm University Press.

Warren, B. A., 1972: Insensitivity of subtropical mode water characteristics to meteorological fluctuations. *Deep Sea Research and Oceanographic Abstracts*, **19 (1)**, 1–19, https://doi.org/10.1016/0011-7471(72)90020-7.

Waterman, S., and B. J. Hoskins, 2013: Eddy shape, orientation, propagation, and mean flow feedback in western boundary current jets. *Journal of Physical Oceanography*, **43 (8)**, 1666–1690, https://doi.org/10.1175/JPO-D-12-0152.1.

Worthington, L. V., 1958: The 18° water in the Sargasso Sea. *Deep Sea Research*, **5 (2)**, 297–305.

Xu, L., J. Liu, Y. Ding, H. Sasaki, H. Cao, J. Dong, and L. Xing, 2025: Submesoscale Effects on Formation of the North Pacific Subtropical Mode Water. *Journal of Physical Oceanography*, **55 (5)**, 559–572, https://doi.org/10.1175/JPO-D-24-0125.1.

Young, W. R., 2012: An Exact Thickness-Weighted Average Formulation of the Boussinesq Equations. *Journal of Physical Oceanography*, **42 (5)**, 692–707, https://doi.org/10.1175/JPO-D-11-0102.1.