Symmetric instability in cross-equatorial western boundary currents

Fraser W. Goldsworth* † and David P. Marshall

Department of Physics, University of Oxford, Oxford, United Kingdom

Helen L. Johnson

Department of Earth Sciences, University of Oxford, Oxford, United Kingdom

*Corresponding author: Fraser W. Goldsworth, frasergocean@gmail.com

†Oxford NERC Environmental Research DTP

ABSTRACT

The upper limb of the Atlantic Meridional Overturning Circulation draws waters with negative potential vorticity from the southern hemisphere into the northern hemisphere. The North Brazil Current is one of the cross-equatorial pathways in which this occurs: upon crossing the equator, fluid parcels must modify their potential vorticity to render them stable to symmetric instability and to merge smoothly with the ocean interior. In this work a linear stability analysis is performed on an idealized western boundary current, dynamically similar to the North Brazil Current, to identify features which are indicative of symmetric instability. Simple two-dimensional numerical models are used to verify the results of the stability analysis. The two-dimensional models and linear stability theory show that symmetric instability in meridional flows does not change when the non-traditional component of the Coriolis force is included, unlike in zonal flows. Idealized three-dimensional numerical models show anti-cyclonic barotropic eddies being spun off as the western boundary current crosses the equator. These eddies become symmetrically unstable a few degrees north of the equator, and their PV is set to zero through the action of the instability. The instability is found to have a clear fingerprint in the spatial Fourier transform of the vertical kinetic energy. An analysis of the water mass formation rates suggest that symmetric instability has a minimal effect on water mass transformation in the model calculations; however, this may be the result of unresolved dynamics, such as secondary Kelvin Helmholtz instabilities, which are important in diabatic transformation.

Significance statement. The Atlantic Meridional Overturning Circulation includes an ocean current that transports heat, carbon and other climatically important tracers from the southern hemisphere into the northern hemisphere. Theoretical considerations suggest that this current may become unstable through the so called "symmetric instability" upon crossing the equator.

In this study, a heirarchy of models is used to investigate how symmetric instability might manifest itself if excited in cross-equatorial flows. We find that when the instability is excited, it generates stacked overturning cells which reorganise the current to make it neutrally-stable to symmetric instability. We hypothesise this process could be occurring in the ocean off the coast of Brazil.

1. Introduction

The Atlantic Meridional Overturning Circulation (AMOC) is a climatically important circulation. In its surface limb, it draws water from the southern hemisphere northwards across the equator, redistributing climatically important tracers such as heat and carbon as it does so (Buckley and Marshall 2015). At 26°N the AMOC transports around 1.25 PW of heat northwards (Bryden et al. 2020), which accounts for around 25% of the meridional transport of heat by the global atmosphere and ocean at this latitude (Srokosz et al. 2012). The AMOC's contribution to poleward heat transport warms Western Europe by around 5°C (Jackson et al. 2015).

The North Brazil Current is a surface-intensified cross-equatorial western boundary current that follows the coast of Brazil from around 10–15°S to around 4–8°N. Here, it retroflects and breaks up into anticyclonic eddies called the North Brazil Current rings which continue northwards to the Carribean (Talley et al. 2011; Fonseca et al. 2004). Both the current and its rings are an important pathway in the northward transport of southern hemisphere waters, and form an important component of the AMOC (Bower et al. 2019). The current itself has an annual mean

transport of around 32 Sv in the upper 600 m of the ocean, with typical peak velocities of ~ 80 $\mathrm{cm\,s^{-1}}$ and a width of ~ 100 km (Johns et al. 1998; Schott et al. 1993). At the equator there is a break down in geostrophy due to the vanishing of the Coriolis parameter; however, at latitudes 1° north or south of the equator, the current is geostrophically balanced to leading order, whilst the rings are in cyclogeostrophic balance (Vianna and de Menezes 2003; Castelão and Johns 2011). The flows considered in the following study are dynamically similar to the North Brazil current; however, the conclusions of this work are expected to apply to cross-equatorial flows in upper-ocean western boundary currents more generally.

The change in sign of the locally vertical component of planetary vorticity at the equator is an important constraint on the way in which water is able to cross from one hemisphere to the other. The Ertel potential vorticity (PV) of a Boussinesq fluid is defined as

$$Q = (\mathbf{f} + \boldsymbol{\nabla} \times \mathbf{u}) \cdot \boldsymbol{\nabla} b\,, \tag{1}$$

where $\mathbf{f}$ is the planetary vorticity, $\mathbf{u}$ is the velocity relative to the Earth's surface, and b is the buoyancy. The PV of a fluid parcel is materially conserved in the absence of mechanical and buoyancy forcing, and neglecting non-linearities in the equation of state. Waters starting in the southern hemisphere typically have negative PV as a result of the vertical component of their planetary vorticity. Close to the equator, the planetary vorticity varies approximately linearly with latitude — the β-plane approximation. As southern hemisphere waters flow northwards across the equator, their planetary vorticity increases and so, to conserve PV, the flow generates anti-cyclonic relative vorticity. Killworth (1991) shows that this requirement to conserve PV inhibits the penetration of fluid from one hemisphere to another, further than a few Rossby deformation radii.

At the equator, the meridional component of the planetary vorticity vector is a maximum. In studies of geophysical fluid dynamics this term is often neglected, which we describe as making the 'traditional Coriolis approximation'. Conversely retention of the meridional component of the planetary vorticity vector is described as including the 'complete', 'full', or 'non-traditional' components of the Coriolis force (Stewart and Dellar 2011). When treating the vertical component of the planetary vorticity as constant, we make the f-plane approximation. Close to the equator we may also make the non-traditional f-plane approximation in which we treat the meridional component of planetary vorticity as constant instead.

In a numerical model of cross-equatorial flow, Edwards and Pedlosky (1998a) find an anti-cyclonic eddy field is generated at the equator. They show that this field is responsible for advecting PV, of the opposite sign to the planetary vorticity, into a viscous boundary layer where it can be dissipated. This provides a mechanism for the modification of PV, whose conservation would otherwise inhibit cross-equatorial flow. The eddy field is observed in other models of cross-equatorial flow (e.g. Jochum and Malanotte-Rizzoli 2003; Goes et al. 2009). In a follow up study, Edwards and Pedlosky (1998b) show that the eddy field is generated by barotropic instability, and is enhanced as the Rossby deformation radius is maximal at the equator. It has been proposed this mechanism is behind the generation of the North Brazil Current rings (Jochum and Malanotte-Rizzoli 2003). The North Brazil Current, however, is a surface current and so viscous bounday layers can play only a limited role, if any, in the modification of PV. As such, turbulent processes in the current's interior are more likely to be of importance. These turbulent processes are of particular interest as they may lead to a short circuiting of the AMOC through enhanced diapycnal mixing. Symmetric instability is one turbulent process that may be capable of modifying the PV in cross-equatorial flows.

Symmetric instability may occur when the PV and the (locally vertical component of the) planetary vorticity of a fluid parcel have opposing signs (Hoskins 1974; Stone 1966)[1]. The change in sign of planetary vorticity at the equator, provided that potential vorticity is conserved, means that a fluid parcel which is initially stable to symmetric instability in one hemisphere, will become symmetrically unstable upon being advected across the equator. Symmetric instability is excited as a result of an imbalance between pressure gradient forces and the Coriolis force, along isopycnals (AMS 2020). The 'symmetric' in symmetric instability comes from the fact that the simplest flows in which it can occur are symmetric in the along stream direction. Flows with $Q \sim 0$ are described as having marginal or neutral PV, and PV of the opposite sign to the planetary vorticity is described as being anomalous. The excitement of symmetric instability in regions of anomalous PV sets up stacked overturning cells which advect PV around to produce a symmetrically stable configuration. In the limit of a flow free of vertical shear (or a flow with horizontal isopycnals), inertial instability is said to occur. Symmetric instability can be thought of as a generalised form of inertial instability along isopycnals (Haine and Marshall 1998; Holton and Hakim 2013). Hua et al. (1997) suggest that the near neutrality of the time mean PV in the interior of the equatorial Pacific may be a result of the excitement of inertial instability in the predominantly zonal flow. In the Atlantic, Castelão and Johns (2011) come to a similar conclusion upon observing marginal PV in the North Brazil Current rings.

[1]It should be noted that the definition of symmetric instability used here is that defined in the AMS glossary of meteorology (2020). Under this 'classical' definition the instability criterion is that $fQ < 0$ (Hoskins 1974), and inertial and gravitational instabilities are both special cases of symmetric instability. This definition is different to that first employed by Thomas et al. (2013) who attempt to characterise symmetric instability in terms of its energetics. Under their taxonomic system the instabilities studied in this work would be described as either inertial or inertial-symmetric. In the literature of the planetary and atmosphere sciences, studies of cross-equatorial flow use the 'classical' definition of symmetric instability (e.g. Joshi 1994; Rodwell and Hoskins 1995). For consistency with this literature base we adopt the same definition here.

It is our hypothesis that symmetric instability is excited in cross-equatorial western boundary currents, such as the North Brazil Current, and that its excitement allows the mixing of water with PV of both signs resulting in a marginally stable PV configuration. The complete Coriolis force has been found to alter the evolution of symmetrically unstable flows (Zeitlin 2018). It is not clear what role the complete Coriolis force may play in the PV modification process. Symmetric instability could also be at play in deep western boundary currents. However, bottom topography may also be important in modifying anomalous PV, and may limit the applicability of the work presented here to deep cross-equatorial flows.

The manuscript starts with a linear stability analysis in section 2, in which the structure of symmetric instability in an idealized western boundary current is derived from theoretical considerations. In section 3 a two-dimensional numerical model is used to investigate the properties of a surface intensified symmetrically unstable western boundary current. Section 4 describes findings from a three-dimensional model of idealized cross-equatorial western boundary currents. Finally, the key findings of this work are summarized in section 5.

2. Linear stability analysis

Hoskins (1974) shows that an inviscid meridional jet, initially in thermal wind balance and symmetric about the meridional axis, may be linearly unstable and produce overturning in the x-z plane. The overturning can be represented as a streamfunction, ψ, where the zonal and vertical velocities are given by $u = -\partial_z \psi$ and $w = \partial_x \psi$ respectively. Hoskins (1974) shows that the streamfunction, to terms linear in ψ, satisfies the partial differential equation

$$\frac{\partial^2}{\partial t^2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right)\psi + \left(N^2\frac{\partial^2}{\partial x^2} - 2f\frac{\partial V}{\partial z}\frac{\partial^2}{\partial x \partial z} + f\zeta\frac{\partial^2}{\partial z^2}\right)\psi = 0\,. \tag{2}$$

Here N is the buoyancy frequency which is assumed to be constant, f is the planetary vorticity, V is the basic state meridional velocity, and ζ is the absolute vorticity of the basic state about the vertical. The equation is easily generalized to flows with a harmonic vertical viscosity, by replacing ∂_t with $\partial_t - A_r \partial^2_{zz}$, where A_r is the vertical viscosity.

Solving for ψ in the viscous case is a difficult problem; however, much can be gained by considering a basic flow that is both barotropic and meridional. Thus the term proportional to the vertical shear of the meridional flow is set to zero. The resulting equation is

$$\left(\frac{\partial}{\partial t} - A_r \frac{\partial^2}{\partial z^2}\right)^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right)\psi + \left(N^2 \frac{\partial^2}{\partial x^2} + f\zeta \frac{\partial^2}{\partial z^2}\right)\psi = 0\,, \tag{3}$$

which, strictly speaking, describes the evolution of an inertial instability due to the flow being free of vertical shear.

We can now try to find solutions of the form $\psi(x,z,t) = \hat{\psi}(x) e^{i(mz-\omega t)}$. Substituting this into (3), we obtain the following boundary value problem:

$$\frac{(\hat{\omega}^2 - N^2)}{m^2} \frac{\mathrm{d}^2\hat{\psi}}{\mathrm{d}x^2} + f\zeta\hat{\psi} = \hat{\omega}^2\hat{\psi}\,, \tag{4}$$

where $\hat{\omega} = \omega + iA_r m^2$. Upon appropriate non-dimensionalisation of the coordinates and variables, (4) is identical to equation 4 of Plougonven and Zeitlin (2009), who identify it as a Schrödinger equation.

From the work of Hoskins (1974), it is known that in the inviscid limit, $\hat{\omega}^2 \sim f^2$. For oceanic western boundary flows, typically $N^2 \gg f^2$. This means we can make the approximation $N^2 - \hat{\omega}^2 \approx N^2$. Plougonven and Zeitlin (2009) show that this is equivalent to making the hydrostatic approximation and that, for a flow similar to the one considered here, there is a negligible effect on the solutions. After making the hydrostatic approximation, we are then left with

$$-\frac{N^2}{m^2}\frac{\mathrm{d}^2\hat{\psi}}{\mathrm{d}x^2} + f\zeta\hat{\psi} \approx \hat{\omega}^2\hat{\psi}\,. \tag{5}$$

The eigenfunctions of the equation are $\hat{\psi}$. They define the horizontal structure of the overturning streamfunction. The eigenvalues of the equation are $\hat{\omega}^2$. If the eigenvalue is negative, then it is possible for ω to be imaginary. If ω is imaginary, then the overturning circulation may either grow or decay exponentially. It is useful to introduce the quantity $\sigma = \mathrm{Im}(\omega)$, which, if positive, corresponds to the exponential growth rate; if negative it gives the decay rate. The value of σ is maximized for the smallest real eigenvalue of equation 5. For each eigenfunction, there exists a spectrum of vertical wave-numbers, each with a characteristic growth rate (or frequency if stable). The relationship between the growth rate and the vertical wavenumber is determined by the eigenvalue, $\hat{\omega}$.

To calculate the eigenfunctions and eigenvalues of the equation, we must first specify a buoyancy frequency and velocity profile, from which vorticity can be calculated. As we are interested in western boundary currents, we will consider an idealized meridional flow: the barotropic Bickley jet. The velocity of the jet can be expressed as

$$V(x) = V_0\left(1 - \tanh^2\left(\frac{x - x_{mid}}{\delta_b}\right)\right) \tag{6}$$

where V_0 is the peak velocity of the jet, x is the across stream coordinate, x_{mid} gives the distance of the peak velocity of the jet from the western boundary and δ_b is the width of the jet. The jet is symmetric in the along stream direction. The jet parameters are set as follows: $V_0 = 0.87$ m s^{-1}, $x_{mid} = 40$ km and $\delta_b = 30$ km. The parameters are chosen to be similar to those used in the two-dimensional and three-dimensional numerical models described in sections 3 and 4 at a depth of 200 m. These parameters are, in turn, loosely based on what is seen in the North Brazil Current (Johns et al. 1998). The velocity profile used is shown in figure 1. In the northern hemisphere, we would expect to see symmetric instability develop in a region to the right of the jet's center. In the southern hemisphere, we may expect to see symmetric instability between the

western boundary and the jet's center. In such a configuration, the western boundary would input anomalous vorticity into the flow, something that this idealized framework is unable to represent. For this reason, and the fact that this study focuses on cross-equatorial flows, we do not apply the linear stability analysis to northward flowing jets in the southern hemisphere.

The buoyancy frequency is set to a value of 5×10^{-3} s^{-1}, which is the mean buoyancy frequency at a depth of between 200 m and 400 m, as estimated from 82 neutral density density profiles taken by an Argo float off the coast of Brazil between January 2016 and February 2017 (Argo 2000). The trajectory of the float and the mean neutral density profile are shown in figure 2. The use of neutral density in calculating the buoyancy frequency means the results of the linear stability analysis will be more applicable to what is seen in the models presented in the following sections rather than the ocean. This is because the neutral density calculation does not reliably preserve vertical buoyancy gradients (Eden and Willebrand 1999). Data from a single Argo float was used as it was readily available and provides a plausible estimate of the density structure of the region. The value of f is set to 1.01×10^{-5} s^{-1}, corresponding to a latitude of approximately 4°N.

For each eigenfunction, which physically corresponds to the horizontal structure of the overturning cell, we can plot the growth rate, σ, as a function of vertical wavenumber, m, and vertical viscosity, A_r. This is done in figure 3. We find that, for a given vertical viscosity, there is a value of the vertical wavenumber which maximizes the growth rate. As the growth is exponential, within a few e-folding timescales, the vertical mode which maximizes σ will dominate the structure of the instability — assuming non-linear effects have not taken hold before this time. Thus, although a discrete set of horizontal modes and a continuous spectrum of vertical modes may be excited, we may expect a single horizontal and vertical mode to dominate the structure of the instability. However, we can only verify this expectation with the use of a numerical model which takes into account the non-linearities neglected here.

Below a maximum 'critical' viscosity, there is a maximum and minimum vertical wavelength at which unstable modes exist; at higher viscosities, all modes are stable. The maximum wavelength is a result of stratification inhibiting vertical motions. Any mode with a wavelength smaller than the minimum will experience strong viscous damping, rendering the mode stable. The minimum vertical wavelength tends to zero in the inviscid limit. The maximum 'critical' viscosity is found when the maximum wavelength allowed by the stratification and the minimum wavelength allowed by viscosity are equal.

It is not *a priori* clear whether the viscosity we are interested in should be a molecular or turbulent viscosity. If one is looking for signs of symmetric instability in a sufficiently coarse ocean model, then it is the turbulent viscosity that will set the vertical length scale. This makes sense as it is the only viscosity the fluid is 'aware' of. For real fluids, matters become more problematic. Griffiths (2003a) suggests that secondary Kelvin-Helmholtz instabilities form as a result of symmetric instability and play a more dominant role in the vertical scale selection than does viscosity. This means that the findings of this linear-stability analysis apply to the results of numerical models which fail to resolve these secondary instabilities, but the relation to what might be observed in meridional western boundary flows in the ocean is more ambiguous.

For a given vertical viscosity we can also plot the two-dimensional structure of the overturning that the instability generates, as shown in figure 4 for a viscosity of 4×10^{-4} m^2 s^{-1} (this viscosity was chosen as it corresponds to the value used in the numerical models presented in sections 3 & 4). We see that the instabilities generate a stack of alternating overturning cells. Overlain in the figure is the absolute vorticity (solid line), highlighting that the cells are stongly localized to the region of negative potential vorticity. Despite the localization, the overturning is non-zero outside of the region of negative absolute vorticity. This enables the mixing of waters with positive and negative PV, which over time will create a neutrally stable PV configuration.

The mode shown in figure 4 has an e-folding timescale of around 2 days. Seawater has a viscosity of 10^{-6} m^2 s^{-1}, which would lead to overturning cells with a much smaller vertical wavelength of 15 m and a smaller e-folding time scale of approximately 1 day. That the timescale is decreased slightly when using a realistic viscosity suggests that the instability will be at least as efficient at neutralizing anomalous PV in the ocean. Differences between the vertical lengthscales seen here and in other works (e.g. Taylor and Ferrari 2009) can be understood in terms of the dependence on vertical viscosity, and the artificially high 'eddy viscosity' used here.

A more rigorous analysis of the solutions to the boundary value problem described in this section is given in Plougonven and Zeitlin (2009), who consider barotropic zonal shear flows on an f-plane. On an f-plane there is no physical distinction between meridional and zonal flows so their findings are broadly applicable, although the vorticity profiles they use differ from those considered here. There also exists a broad literature base on linear stability analyses of symmetric and inertial instabilities on a β-plane, with a focus on zonal currents (e.g. Kloosterziel et al. 2017; Ribstein et al. 2014; Griffiths 2003b; Hua et al. 1997).

3. Two-dimensional numerical models

The previous section presents results from a linear stability analysis of idealized, meridionally symmetric flows, situated north of the equator at approximately 4°N. This section details a series of numerical simulations in which the symmetric instability of more realistic currents is probed. Modifications include finite depth, variable stratification and surface intensification. The two-dimensional nature of the flow is retained and the β-effect still neglected, via the imposition of periodic meridional boundary conditions on a domain one grid cell thick. This prevents the development of barotropic instability, meaning the evolution of the model will be dominated by symmetric instability.

a. Methods

The numerical simulations are performed using the MITgcm model (Marshall et al. 1997). The model domain consists of a channel 400 km wide in the zonal direction and periodic in the meridional direction. The horizontal resolution is 2 km. In the vertical there are 160 depth levels, varying in size from 6.25 m at the surface to 25 m at the bottom, giving a total depth of 1,500 m.

Simulations are run on an f-plane, with $f = 1.01 \times 10^{-5}$ s^{-1}, corresponding to a latitude of 4°N, with the non-traditional component of the Coriolis force included also, taking a value of $F_{NT} = 1.45 \times 10^{-4}$ s^{-1}. A control integration is also run in which the non-traditional component of the Coriolis force is neglected. A further set of integrations at 40°N and 0°N, both with and without the non-traditional component of the Coriolis force, are performed. For the integration at 40°N, $f = 9.35 \times 10^{-5}$ s^{-1} and $F_{NT} = 1.11 \times 10^{-4}$ s^{-1}; at 0°N, $f = 0$ s^{-1} and $F_{NT} = 1.45 \times 10^{-4}$ s^{-1}. In all cases, no qualitative difference between integrations with and without the complete Coriolis force is found; this finding is discussed at length in section 3c.

Closure of the momentum equations in the horizontal is provided by an adaptive biharmonic Smagorinsky viscosity. This is chosen to minimize damping at the length-scales of interest (Smagorinsky 1963; Griffies and Hallberg 2000), with the choice inspired by its successful use in Brannigan (2016) which attempts to resolve similar sub-mesoscale processes. A Laplacian vertical viscosity, A_r, is used and set to a constant value of 4×10^{-4} m^2 s^{-1} for all integrations, apart from a viscous integration for which a value of 6×10^{-3} m^2 s^{-1} is used. The standard value is chosen to ensure the vertical structure of the symmetric instability can be adequately resolved by the model grid.

A linear equation of state is employed:

$$\rho = \rho_0 \big(1 - \alpha_T (T - T_0)\big), \tag{7}$$

where ρ_0 is the background density, T is temperature, T_0 is a reference temperature, and α_T is the thermal expansion coefficient. The absence of salinity means that changes in density are modulated solely by changes in temperature. The value of ρ_0 is set to 1023.35 kg m^{-3}, α_T set to 2×10^{-4} K^{-1}, and T_0 set to 30°C. The linear equation of state is used to avoid the complexities added by non-linear effects. The diffusivity of temperature is set to 1×10^{-5} m^2 s^{-1}. The temperature advection scheme used is the second order-moment Prather with limiter (Prather 1986). The initial density profile is based on observations taken by an Argo float in the tropical Atlantic (Argo 2000). The density profile and float trajectory is shown in figure 2, with the temperature profile in the model chosen to match the observed density profile.

The initial velocity profile is based on that of the surface intensified Bickley jet. This can be expressed as

$$V(x,z) = V_0\left(1-\tanh^2\left(\frac{x-x_{mid}}{\delta_b}\right)\right)\frac{z+H}{H}, \tag{8}$$

where z is the vertical coordinate which becomes increasingly negative below the surface, and H is the absolute value of the depth. The jet parameters are set as follows: $V_0 = 1$ m s^{-1}, $x_{mid} = 40$ km, $\delta_b = 30$ km, $H = 1,500$ m. The profile is shown in figure 1b. To test the sensitivity of the model to the jet's position, two control simulations are also run with x_{mid} set to either 20 km or 80 km: the resulting instability shows no qualitative dependence on the jet's position. The jet parameters are chosen to be similar to those observed in the North Brazil Current by Johns et al. (1998). However, it should be noted that the jet described in this work is more intense ($V_{max} \sim 1$ m s^{-1} rather than ~ 0.9 m s^{-1}), has less vertical shear ($\partial_z V \sim 6.6\times10^{-4}$ m^{-1} s^{-1} rather than $\sim 3.5\times10^{-3}$ m^{-1} s^{-1}), and is deeper ($H \sim 1,500$ m rather than ~ 800 m) than that observed by Johns et al. (1998), which puts some limitations on the direct applicability of the results presented here to the North Brazil Current. The most significant differences occur in the shear and depth. The depth is chosen to

be large, so as to prevent the bottom boundary layer from having too strong an influence on the evolution of any instabilities.

In the 4°N integrations, the jet prescribed is symmetrically unstable from the outset. This is intentional as we wish to see how a symmetric instability evolves in a 'plausible' western boundary current. In a two-dimensional numerical model, the meridional advection of southern hemisphere waters with negative PV into the domain is not possible. As such we must initialise the numerical integration with waters of negative PV already in place.

At the surface boundary a rigid lid condition is used. The lateral boundary condition at the eastern and western edges of the domain is set to no-slip. At the bottom, a free-slip boundary condition is used. The sensitivity of the simulations to the choice of boundary condition has been tested by performing model integrations with free-surface, free-slip lateral, and no-slip bottom boundary conditions: in each case, no qualitative differences in the resulting instability are observed.

b. Results and their relation to the linear stability analysis

Snapshots of PV at a range of times from the standard two-dimensional numerical integration are shown in figure 5. We see that after around 5 days the PV distribution begins to change. The timescale (the inverse of the growth rate, $1/\sigma$) predicted by the linear stability analysis is ~ 2 days, which is of the right order of magnitude for what is seen here. In the final panel, we can see how the PV in the initially unstable region has been modified and set to a state of marginal stability — i.e. $Q = 0$. This modification of PV is driven by the underlying overturning motions predicted by the linear stability analysis. Due to the along stream symmetry, we can rearrange the equation for the conservation of PV to give

$$\frac{\partial Q}{\partial t} = \frac{\partial \psi}{\partial z}\frac{\partial Q}{\partial x} - \frac{\partial \psi}{\partial x}\frac{\partial Q}{\partial z}, \tag{9}$$

which makes the link between the unstable overturning generated by the instability and the redistribution of PV explicit[2].

Figure 6 shows the overturning streamfunction generated in the standard two-dimensional numerical integration. We see structures similar to those predicted by the linear stability analysis (c.f. figure 4). The vertical Fourier transform of the streamfunction is taken at two weeks. At the longitude of the absolute vorticity minimum, there is a maximum in the power spectrum with a vertical wavelength of 100 m. This estimate of the size of the overturning cells is remarkably similar to the prediction of 105 m made by the linear stability analysis. We would expect fairly close agreement as the parameters used in the linear stability analysis are based on those found between 200 m and 400 m of the jet used in this model.

The resolution of the model means that secondary Kelvin-Helmholtz instabilities that typically accompany the excitement of symmetric instability are absent. The resolution and aspect ratio of the model may also be insufficient to fully resolve the dynamics of the symmetric instability if the overturning cells become aligned with isopycnals as expected (Bachman and Taylor 2014). Furthermore, the use of an artificially high vertical viscosity, to ensure numerical stability, means the growth rate of the instability is suppressed and its vertical extent exaggerated. Although symmetric instability may look different in the ocean, the conclusion that symmetric instability is efficient at eliminating anomalous PV within a few degrees of the equator seems unlikely to change.

The linear stability analysis suggests that by setting the viscosity to be sufficiently high, the unstable motions can be suppressed. Integrations with a high viscosity of 6×10^{-3} m^2 s^{-1} are performed and, indeed, no instability develops. The linear stability analysis also predicts no

[2]This equation neglects the frictional dissipation of PV; however, this contribution is generally much smaller than the contribution from advection.

instability if the experiment is performed at a sufficiently high latitude, or at the equator: model integrations at 40°N and 0°N verify this is indeed the case (not shown).

c. Dependence of the instability on the inclusion of the complete Coriolis Force

Zeitlin (2018) states that "symmetric instability drastically changes upon inclusion of the full Coriolis force." The author goes on to perform a linear stability analysis of a zonally symmetric shear flow on an f-plane, taking into account the non-traditional component of the full Coriolis force. They find that the growth rate and spatial structure of unstable modes may be significantly altered by inclusion of the full Coriolis force. Naïvely, one may expect to observe drastic changes in the above numerical simulations, given that on an ordinary f-plane there is no difference between a meridional and zonal flow (as the system has no preferred orientation). However, the inclusion of the non-traditional component of the Coriolis force breaks this symmetry and reintroduces a natural meridional direction. To understand how symmetric instability in meridional flows changes under the complete Coriolis force, we must repeat the linear stability analysis of Zeitlin (2018) but for the case of a meridional flow.

The Boussinesq equations of motion for a meridionally symmetric fluid on an f-plane, with the complete Coriolis force, and in the limit in which the horizontal length scale is much larger than the vertical length scale, are

$$\frac{Du}{Dt} - fv + F_{NT}w - \frac{\partial \phi}{\partial x} = A_r \frac{\partial^2}{\partial z^2} u\,, \tag{10a}$$

$$\frac{Dv}{Dt} + fu = A_r \frac{\partial^2}{\partial z^2} v\,, \tag{10b}$$

$$\frac{Dw}{Dt} - F_{NT}u - \frac{\partial \phi}{\partial z} - b = A_r \frac{\partial^2}{\partial z^2} w\,, \tag{10c}$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0\,, \tag{10d}$$

and

$$\frac{Db}{Dt} = \kappa \frac{\partial^2}{\partial z^2} b \,. \tag{10e}$$

Here, ϕ is the geopotential pressure, κ is the buoyancy diffusion coefficient and F_{NT} is the non-traditional component of the Coriolis parameter.

We now consider a flow which is, initially, purely meridional, with $(u, v, w) = (0, V, 0)$, $\phi = \Phi$ and $b = B$. The flow is in geostrophic balance, hydrostatically balanced and in equilibrium. Mathematically

$$fV = \frac{\partial \Phi}{\partial x}, \tag{11a}$$

$$B = \frac{\partial \Phi}{\partial z}, \tag{11b}$$

$$\frac{\partial^2 V}{\partial z^2} = 0, \tag{11c}$$

and

$$\frac{\partial^2 B}{\partial z^2} = 0 \,. \tag{11d}$$

We can then perturb this balanced flow, giving perturbed variables $(u, v, w) = (u', V + v', w')$, $\phi = \Phi + \phi'$ and $b = B + b'$. Substituting these values into the equations of motion and considering terms only of linear order or lower in the perturbed variables, we find that

$$\frac{\partial u'}{\partial t'} + F_{NT} w' - f v' + \frac{\partial \phi'}{\partial x} = 0, \tag{12a}$$

$$\frac{\partial v'}{\partial t'} + u' \frac{\partial V}{\partial x} + w' \frac{\partial V}{\partial z} + f u' = 0, \tag{12b}$$

$$\frac{\partial w'}{\partial t'} + \frac{\partial \phi'}{\partial z} - b' - F_{NT} u' = 0, \tag{12c}$$

$$\frac{\partial b'}{\partial t} - \kappa \frac{\partial^2 b'}{\partial z^2} + w' \frac{\partial B}{\partial z} + u' \frac{\partial B}{\partial x} = 0, \tag{12d}$$

and

$$\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z} = 0, \tag{12e}$$

where the operator $\partial/\partial t'$ is defined as

$$\frac{\partial}{\partial t'} = \frac{\partial}{\partial t} - A_r \frac{\partial^2}{\partial z^2}\,. \tag{13}$$

From equation (12e) we see that, as in the linear stability analysis (section 2), we can write u' and w' in terms of an overturning streamfunction ψ, where $u' = -\partial_z \psi$ and $w' = \partial_x \psi$. We now obtain an equation of motion for the overturning streamfunction. The first step is to differentiate the horizontal and vertical momentum equations (equations 12a and 12c) with respect to the vertical and zonal coordinates respectively. Subtracting the two equations gives

$$-\frac{\partial}{\partial t'}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right)\psi - f\frac{\partial v'}{\partial z} + \frac{\partial b'}{\partial x} + F_{NT}\left(\frac{\partial u'}{\partial x} + \frac{\partial w'}{\partial z}\right) = 0\,. \tag{14}$$

The last term on the left hand side is the only term containing an explicit dependence on F_{NT}, and we know from equation (12e) that it is equal to zero. The only other way the equation of motion for the streamfunction could gain a dependence on F_{NT} is through the evolution of v' or b'. b' is a function of the streamfunction and the initial buoyancy field and v' depends on the streamfunction and the initial meridional velocity profile. Thus, the overturning driving the redistribution of PV evolves in exactly the same manner regardless of whether there is rotation about the meridional axis or not.

The PV itself is also independent of the complete Coriolis force for a meridionally symmetric flow. This can be seen by explicitly evaluating equation (1), giving

$$Q = \left(f + \frac{\partial v}{\partial x}\right)\frac{\partial b}{\partial z} - \frac{\partial v}{\partial z}\frac{\partial b}{\partial x}\,. \tag{15}$$

From equation (9) we can see that as both the streamfunction and PV are independent of F_{NT}, the evolution of the symmetric instability will not depend on it either. This finding is not in any way contradictory to the findings of Zeitlin (2018). The difference arises due to the asymmetry between the purely meridional flow considered here and the zonal flow considered in the aforementioned study.

It is, in fact, possible to make a more general statement about the types of forces which leave the overturning unchanged. We can modify the momentum equations (12a and 12c) with the addition of any irrotational force acting in the xz-plane and which satisfies the relationship

$$\frac{\partial \mathcal{F}_x}{\partial z} + \frac{\partial \mathcal{F}_z}{\partial x} = 0, \tag{16}$$

where $\mathcal{F}_x$ and $\mathcal{F}_z$ are the zonal and meridional components of the force respectively. This can be understood in the context of Marshall and Pillar (2011) as follows — an irrotational force is divergent and so will project on to the pressure gradient terms of the momentum equation. A rotational force is non-divergent and so projects entirely onto the acceleration term. An irrotational (divergent) force is not able to alter the acceleration term. In the system described above, the inclusion of the complete Coriolis force may alter the pressure field but not the motions within the xz-plane due to its irrotational nature.

Although the crossing of the equator is a meridional phenomenon, western boundary currents, such as the North Brazil Current, will be oriented at some angle to a meridian, having both zonal and meridional components of velocity. In the zonal limit, symmetric instability can change drastically with the inclusion of the full Coriolis force, whereas in the meridional limit there is no change at all. For a realistic (not purely meridional) western boundary current crossing the equator, the structure of symmetric instability may therefore have some dependence on the complete Coriolis force.

The relative importance of non-traditional effects depends on the direction of the current relative to the meridional direction. For a current oriented at an angle θ to the meridional direction, the findings of Zeitlin (2018) apply but with the value of F_{NT} scaling with $\sin\theta$. Zeitlin (2018) defines a "non-traditionality" parameter and using the findings of our work we can generalise it to flows

with a meridional component giving

$$\gamma = \frac{\cot\phi \sin\theta H}{L} \tag{17}$$

where γ is the "non-traditionality parameter" and ϕ is the latitude. Non-traditional effects are important when $\gamma \sim 1$. Close to the equator, the coast of Brazil forms an angle of $\theta \sim 60°$ to the meridian. Using this value of θ along with $\phi = 4°$N, $H = 100$ m and $L = 30$ km gives $\gamma = 0.04$, suggesting that non-traditional effects are unlikely to be hugely important off the coast of Brazil.

4. Three-dimensional numerical models

Although a useful tool in understanding the processes at play in cross-equatorial flows, the two-dimensional numerical model fails to capture key aspects of the dynamics. Perhaps the most important are the absence of the β-effect and the supression of meridionally asymmetric motions, both of which arise from the imposition of along-stream symmetry. It is the variation and change in sign of planetary vorticity at the equator which we suspect will lead to symmetric instability. Furthermore, Edwards and Pedlosky (1998b) show that barotropic eddies are an important and robust feature of cross-equatorial flows. These motions cannot form if there is no meridional variation in the flow. To fully understand what is going on in the ocean, we must relax the requirement of along-stream symmetry and turn to three-dimensional numerical models. This section describes an idealized three-dimensional model of the tropical Atlantic based on the two-dimensional model discussed previously.

a. Methods

The three-dimensional model uses a configuration based on the two-dimensional model described in section 3a. The differences are outlined below.

The horizontal domain is 816 km in the zonal and 2,688 km in the meridional direction. The horizontal resolution is 2 km. The southernmost boundary is located 512 km south of the equator and the northern boundary 2,176 km to the north. The velocity at the northern and southern boundaries is prescribed and takes the form of the surface intensified Bickley jet, as shown in figure 1b. The zonal velocity is initially set to zero and the meridional velocity to the same surface intensified Bickley jet prescribed at the boundaries. As in the models previously discussed, this means that initial PV configuration is unstable to symmetric instabilities. In the northern hemisphere, within around 3 weeks, the majority of the negative PV initially present has been neutralized by an initial flurry of symmetric instability (in the standard viscosity runs). After this time the largest source of negative PV in the northern hemisphere is waters advected across the equator — this can be seen clearly in the animations included in the supplementary materials.

To absorb any incoming waves or eddies, there exists a sponge region within the domain, which relaxes the velocity to the Bickley jet profile prescribed at the boundaries. The sponges stretch from 350 km south of the northern and 40 km north of the southern domain boundaries. The inverse relaxation timescale in these sponge regions varies according to a tanh function from 0 s^{-1} in the interior of the sponge to 2×10^{-5} s^{-1} at the open boundary. The sponge regions are not shown in any figures.

Unlike in the two-dimensional model integrations, the three-dimensional integrations evolve in qualitatively different ways depending on whether a no-slip or free-slip lateral boundary condition is imposed. The no-slip boundary condition is taken to be the standard choice, although results from free-slip integrations are also presented. A time step of 144 seconds is used for the no-slip model and 72 seconds for the free-slip model, which has a higher Reynolds number.

Integrations are performed with either a standard vertical viscosity of 4×10^{-4} $m^2\,s^{-1}$, or a high vertical viscosity of 1×10^{-2} $m^2\,s^{-1}$. The high viscosity is larger than the value of 6×10^{-3} $m^2\,s^{-1}$

used in the two-dimensional high viscosity runs. This is because the original value was insufficient to shut down the excitement of symmetric instability at high latitudes. The larger value is still not strong enough to completely shut down the excitement of symmetric instability; however, it does appreciably suppress the growth of the instability.

The β-plane approximation is made, which provides a linear latitudinal variation of planetary vorticity of 2.3×10^{-11} m^{-1} s^{-1}. The non-traditional component of the Coriolis force is also included by setting F_{NT} to a constant value[3] of 1.5×10^{-4} s^{-1}.

The same thermodynamic scheme and initial stratification is used as in the two-dimensional numerical models previously discussed. In much of the following discussion analyses are performed at a depth of 50 m. This depth is chosen as it sits below the models mixed-layer but is still close to the surface where the excitement of symmetric instability is expected to be most vigorous.

b. Results & discussion

1) No-slip lateral boundary conditions

Figure 7 shows snapshots of the PV field in the standard three-dimensional no-slip integration, at a depth of 50 m. Close to the equator we see the spinning up of anticyclonic eddies of up to 200 km in diameter. The eddies are dynamically similar to those seen in other studies of cross-equatorial flow and are a result of barotropic instability (e.g. Edwards and Pedlosky 1998a,b; Jochum and Malanotte-Rizzoli 2003; Goes et al. 2009). Between around 250 km and 1000 km north of the equator, smaller scale features within the eddies are visible: these arise from the excitement of symmetric instability.

[3]Strictly speaking, in a spherical geometry we would expect $F_{NT} = 2\Omega \cos\theta$ where θ is the latitude. At the equator, the variation in F_{NT} with latitude is minimal, and over the domain considered here, F_{NT} is constant to 2 significant figures. As such it was decided to ignore the meridional variation in F_{NT} and make the "non-traditional f -plane" approximation.

There is an interesting interplay between the barotropic and symmetric instabilities. We see in the northern hemisphere of figure 7a that the eddy with anticyclonic PV initially penetrates a few hundred kilometers, before it begins to retroflect. Within around a week the PV within the eddy has been reduced and it is able to propagate further northwards (see figure 7b). If we look to the sub-mesoscale, we see features with large spatial PV gradients over short distances. This is a result of the excitement of symmetric instability which acts to redistribute PV within the eddy. By reducing its PV the eddy is able to propagate even further north (figure 7c). As the eddy moves further northwards the growth rate of symmetric instability increases and any remaining regions of negative PV become increasingly unstable. We are left with an anticyclonic eddy with approximately neutral PV, corresponding to a state of marginal stability. Figure 8 shows a comparison of the relative and potential vorticity of an eddy which has undergone this process. We see that close to the equator regions of negative PV correspond to regions of negative relative vorticity. In a region from around 500 km to 800 km north of the equator the relationship is less clear. North of this region, the planetary vorticity is larger than the relative vorticity and contributes the most to the absolute vorticity. From figure 7 (and more clearly its animated counterpart, figure 7S, in the supplementary materials), the excitement of symmetric instability is what seems to be driving the transition from one regime to the other. South of a distance 500 km north of the equator, the relative vorticity is negative. Between 500 km and 800 km north, symmetric instability is excited. Waters leaving this region then have approximately neutral PV as a result of the excitement of symmetric instability.

The offset of this symmetrically unstable region from the equator can be understood in terms of the growth rate of the instability.

Firstly, in the inviscid limit the square of the growth rate of symmetric instability is proportional to $f(f+\xi)$, where ξ is the relative vorticity of the fluid. To first order (and in the absence of symmetric instability), changes in the planetary vorticity are approximately balanced by the

relative vorticity, meaning $(f+\xi)$ is approximately constant along flowlines. This gives a growth rate that grows linearly with f and hence latitude. Close to the equator where f is small, the growth rate will be so small that other processes will dominate over symmetric instability. When a fluid parcel is advected further north, its growth rate will increase until the excitement of symmetric instability becomes apparent. At this point, symmetric instability begins to modify the PV and absolute vorticity, removing anomalous vorticity from the flow. At higher latitudes still, the source of negative PV that the instability feeds off has been largely depleted. This means we only see symmetric instability occurring within a finite latitude range.

Secondly, the fluid takes a finite time to show signs of symmetric instability. During this time the fluid will have been advected northwards away from the equator. This distance, d, can be crudely estimated by $d = V\tau$, where V is a typical velocity of the fluid and τ is the timescale it takes for symmetric instability to become apparent in the flow. Taking $V \sim 80$ cm s^{-1} and $\tau \sim 4$ days gives $d \sim 300$ km. From this, we shouldn't expect the instability to be well developed either at or close to the equator.

It is interesting to look at snapshots of the PV from the three-dimensional model at a fixed latitude, to track how the symmetric instability evolves over time, akin to what is shown in figure 5. This is done in figure 9. Initially the instability behaves similarly to the two-dimensional case, although the features are slightly distorted by 'noise' from passing eddies and inertial waves. In more realistic model set ups, and in the ocean, it may not be possible to see the initial, clean evolution of symmetric instability. As such, other diagnostic quantities may be required to detect its excitement.

As with the two-dimensional numerical model, the coarseness of the grid and use of an enhanced eddy viscosity are key limitations of the three-dimensional integrations presented here. The grid is unable to resolve fully the evolution of secondary instabilities, may not correctly resolve motion

along slanted isopycnals, will suppress the growth rate, and will enhance the vertical extent of the instability. As before, it seems unlikely that this will change the fact that symmetric instability is efficient at neutralising anomalous PV in waters that have crossed the equator, especially due to the larger growth rate we expect to see when a realistic viscosity is used.

The PV evolution of the viscous no-slip integration is shown in figure 10. The results appear qualitatively similar to those of Edwards and Pedlosky (1998a) whose model does not permit symmetric instability. In this model, lateral friction is leading to dissipation of PV, which is not as dominant in the standard no-slip integration. At some latitudes, sub-mesoscale patterns are starting to become apparent, particularly within the anticyclonic eddies. Figure 11 shows how the PV varies with depth. Immediately clear is the presence of symmetric instability like features, suggesting the high viscosity has only partially suppressed symmetric instability. Note how the vertical length scale of the disturbance is larger in this integration than in the standard no-slip one. This can be understood with reference to figure 3 — higher viscosities lead to larger vertical length scales.

2) Free-slip lateral boundary conditions

The behaviour of the three-dimensional model under no-slip and free-slip lateral boundary conditions is qualitatively different. The main difference is the behavior of the barotropic eddies. The sensitivity to the choice of lateral boundary condition is not seen in the two-dimensional models as they do not develop these barotropic eddies. The differences can be seen by comparing figures 7 and 12, and more clearly in their animated counterparts, figures 7S and 12S of the supplementary materials.

In the free-slip integrations, the eddies grow to be larger, and propagate northwards more quickly, than in the no-slip integrations. They do not retroflect in the same way as before. As they

propagate northwards, they entrain waters of negative PV and elongate, creating a concentrated pool of symmetrically unstable waters. When the eddies do become unstable, the process is much more explosive. This can be seen in figure 12 and its animated counterpart, figure 12S, in the supplementary materials. As the eddies propagate northwards more quickly, the symmetrically unstable region appears to be shifted northwards. This is because it takes a similar time for symmetric instability to develop, but the eddies have moved further north of the equator during this time.

The PV evolution from a free-slip viscous integration is shown in figure 14, reinforcing that, with the noise of symmetric instability removed, the barotropic eddies behave very differently under the two different lateral boundary conditions. This raises the question of what the most appropriate boundary condition is. For the idealized model setup, this is likely the no-slip lateral condition. In a realistic model, given the small aspect ratio of the problem, a no-slip bottom boundary and free-slip lateral boundary condition, along with variable bathymetry might better represent the physics. This is a result of gently sloping ocean bathymetry in the region meaning that the sea floor is better approximated as a bottom, rather than lateral, boundary.

3) Correlations between relative and potential vorticity

It is useful to be able to define concretely the latitudes at which symmetric instability is occurring. Thus far, this has been done visually by looking at the regions in which sub-mesoscale patterns begin to form in areas of anticyclonic vorticity. It has been noted that regions in which symmetric instability is occurring mark a transition from a regime in which PV is dominated by relative vorticity to one in which the planetary vorticity dominates. Figure 8 suggests it may be possible to identify latitudes at which symmetric instability is occurring by considering correlations between relative vorticity and PV.

The relative vorticity and PV are interpolated onto the same grid. The Pearson correlation coefficient, r, of the vertical component of relative vorticity and PV between 21 days and 49 days is calculated for all grid points between the western boundary and 400 km east at 50 m depth for each latitude. The calculation starts at 21 days so that any instability resulting from the initial conditions has died down. How the correlation coefficient varies with latitude is shown in figure 15, for both the no-slip and free-slip, and standard and viscous, integrations.

In figure 15, we see that the vertical component of relative vorticity and PV are very strongly correlated in the southern hemisphere. At around 250 km north of the equator, there is an abrupt change in the correlation for the standard no-slip and free-slip integrations. The latitude of the change seems to correspond to the latitude north of which we observe symmetric instability. In the case of the no-slip integration, r reduces to around 0.8, and for the free-slip integration it is reduced even further, to less than 0.6. This makes sense — in the free-slip experiment the excitement of symmetric instability appears qualitatively more vigorous than in the no-slip experiment.

We see a much smaller and more gradual drop in the correlation for the viscous integrations, and the decrease in r appears to begin at much higher latitudes — around 600 km for the free-slip and 1,000 km for the no-slip integration. This is consistent with what we see in the models: at low latitudes we see little evidence of symmetric instability; however, its effects become apparent at higher latitudes.

We can understand the change in correlations in terms of the PV of a meridional flow, given by

$$Q \approx \left(N^2 f - M^2 \frac{\partial V}{\partial z} \right) + N^2 \xi \tag{18}$$

where N^2 and M^2 are the vertical and horizontal buoyancy gradients, respectively. In the absence of symmetric instability, at a given latitude, one would expect there to be a strong correlation between Q and ξ. If N^2 and f are approximately constant, and M^2 and $\partial_z V$ are small, then the bracketed

term of the above equation will be approximately constant and a linear relationship between Q and ξ will exist. When symmetric instability is excited, localized vortex stretching and changes in the vertical shear cause a breakdown of this linear relationship. As such, one would expect the correlation between the vertical component of relative vorticity and PV to be much lower. The fact that the correlation changes only slightly in the viscous runs, which exhibit barotropic eddies but suppressed excitement of symmetric instability, suggests that the changes in the linear relationship are the result of symmetric instability.

4) Spectral energy density

Symmetric instability is a sub-mesoscale instability, suggesting that when it is excited, we would expect to see an increase in the amount of energy at small length scales. Yankovsky and Legg (2019) find that within their models of symmetric instability in Arctic overflows, symmetric instability has a fingerprint in the spatial Fourier transform of the vertical kinetic energy (KE). This technique is used here to attempt to diagnose symmetric instability in a similar manner.

For the standard three-dimensional no-slip integration, the quantity $w^2/2$ is calculated at a depth of 50 m from the western boundary to 200 km east, and from 250 km to 750 km north. The power spectrum is taken along the zonal dimension and the meridional mean taken. This is plotted at 1 week intervals in figure 16a. The same is done for the three-dimensional viscous no-slip integration in figure 16b. The same procedure is performed for the single latitude of the two-dimensional standard and viscous no-slip models and the results shown in figures 16c and 16d respectively.

In each of the panels a, b & c of figure 16, we see a flattening of the power spectra over time. This corresponds to the development of vertical motions with a structure that varies over small

length scales. In the case of the standard three-dimensional model, within around four weeks the power spectrum reaches an equilibrium.

In the viscous three-dimensional model it takes around three weeks for the flattening to take place. This is due to the increased vertical viscosity suppressing the growth rate. Interestingly, however, we do still see the flattening effect of symmetric instability at later times, as the viscosity is not sufficient to completely inhibit its development.

In the standard two-dimensional model, after an initial flattening of the spectra there is a subsequent steepening from four weeks onwards. After four weeks, the two-dimensional model has used up most of its finite reserves of negative PV which fuel symmetric instability. This means that there is no process to sustain the small scale vertical motions and they begin to die away. This is in contrast to the three-dimensional model in which the supply of negative PV is constantly being replenished by the advection of southern hemisphere waters across the equator. As the two-dimensional model is not capable of producing barotropic eddies, we can be fairly certain that the changes in the power spectrum are not related to them. The increased variability relative to the three-dimensional models is due to the removal of the meridional averaging step.

In the viscous two-dimensional model, in contrast to its three-dimensional counterpart, we see no flattening at all in the power spectra[4]. This is because the vertical viscosity is sufficient to completely inhibit symmetric instability in this model. The stationarity of the power spectra allows us to say with some confidence that the flattening seen in the other models is due to the presence of symmetric instability. Indeed, the flattening of power spectra may prove useful for identifying symmetric instability in more complex ocean models and in observations, in which the signature of symmetric instability may otherwise be less obvious.

[4]Note that the viscous two-dimensional model uses a viscosity of 6×10^{-3} m^2 s^{-1} compared to 10^{-2} m^2 s^{-1} for the viscous three-dimensional model.

The power spectra appear to show a lot of power concentrated at the grid-scale. Inspection of the vertical velocities in the region suggest this is not the result of grid-scale noise, but rather physical flow structures. Note that the axes in figure 16 are logarithmic and so may appear to exaggerate the amount of energy at the grid-scale.

5) Transformation of water masses

Symmetric instability leads to the generation of small-scale overturning cells. Such cells may be expected to lead to the mixing of waters with different density and contribute to water mass transformation. In the Walin (1982) water mass transformation framework, the rate of formation of water between two isopycnal surfaces ρ and $\rho + \Delta\rho$ is given by

$$\mathcal{F} = \frac{\partial \mathcal{V}}{\partial t} + \Delta\psi = G(\rho) - G(\rho + \Delta\rho) \tag{19}$$

where $\mathcal{F}$ is the rate of formation of fluid between the isopycnals, $\mathcal{V}$ is the volume bounded by the isopycnals, $\Delta\psi$ is the net volume flux out of the region, and $G(\rho)$ is the diapycnal volume flux across the ρ-isopycnal — also known as the water mass transformation. By convention a positive value of G corresponds to a flux from lighter to denser waters. These quantities are shown diagrammatically in figure 17.

The water mass formation is calculated for density classes of $\sigma < 23.45$, $23.45 \leq \sigma < 26.50$ and $\sigma \geq 26.50$, where $\sigma = \rho - 1000$ kg m^{-3}. Physically these correspond to a surface layer, the pycnocline and deep ocean, as shown in figure 2 alongside the initial stratification profile. The formation is diagnosed in latitude bands with boundaries at -500 km, -250 km, 0 km, 250 km, 750 km, and 1500 km for both the standard no-slip and viscous no-slip integrations. There is a large amount of temporal variability in the formation and so it is hard to discern any trends in the volumes of the chosen density layers. To remove this variability, the rolling average is taken over 25 output time-steps (corresponding to approximately 2.9 days), and is shown in figure 18.

By cumulatively integrating the time series, it is easier to identify long term trends in formation. Figure 19 shows the cumulative formation in each of the density classes and latitude bands.

South of 250 km, there are no significant changes in the volume of the surface layer for either the standard or viscous runs. North of this latitude, in both models we see a trend corresponding to a decrease in volume. In the 250 km to 750 km band, the trend is most significant in the standard run; however, in the 750 km to 1,500 km band the volume change of the viscous run is more prominent. These latitudes are where we see the excitement of symmetric instability, which is mixing dense waters from below the surface layer with waters within it. This leads to a reduction of waters in the lightest density class.

The growth rate of symmetric instability in the extra viscous integration is suppressed at lower latitudes. At higher latitudes, where the instability has had longer to develop, its effects can still be seen in the depletion of the mixed layer. The formation is stronger for the viscous integration than the standard in the northernmost latitude band. The supply of negative PV which fuels the instability is lower in the case of the standard integration as it has been neutralized further south. This reduced supply is what causes lower formation in the standard integration than in the viscous integration.

For the pycnocline, in the two northernmost latitude bands, we see differences in the formation between the standard and viscous integrations. For the standard integration, we see both a net formation and depletion of pycnocline waters depending on the time-scales we look at. For the viscous integration, we see a similar time-varying formation-depletion pattern within the pycnocline, however, at later times in the northernmost band, we also see a trend of net depletion of pycnocline waters. Looking to the deep ocean (in the same latitude band), we see a net formation of deep waters, similar in magnitude to the depletion we see in the surface layer and pycnocline waters. This can be understood in terms of the dependence of the vertical scale of symmetric

instability on the viscosity. Figure 3 suggests that as we increase the viscosity the size of the overturning cells will increase. For the viscosity used in the viscous integration we might expect a $\lambda \sim 10^2$ m. This can lead to overturning cells able to connect the deep ocean and the surface layer, as seen in figure 11.

The magnitude of the filtered water mass formation is typically small compared to the meridional transport of the modeled current. The maximum formation rate is around 4.8 Sv; however, more generally, symmetric instability is responsible for instantaneous formation rates of no more than 2 Sv. The cumulatively integrated formation for the standard run suggests that the net effect of symmetric instability on water mass transformation is small, with large instantaneous formation rates balancing over time periods of several weeks.

The water mass formation rates calculated here may have limited applicability to the ocean. Here, we have found that the rates are affected by the vertical viscosity used, suggesting different formation rates would be found if a lower viscosity were used. Yankovsky and Legg (2019) find that mixing by symmetric instability is largely adiabatic, with irreversible mixing resulting from secondary shear instabilities. The model grid used here is too coarse to resolve such secondary instabilities, suggesting formation rates may be higher if a finer grid were used. Similarly, the model resolution may prevent the accurate representation of symmetric instability when the overturning cells are oriented along isopycnals. It seems unlikely that this limitation will alter the key finding that symmetric instability is efficient at removing anomalous PV originating from the opposing hemisphere.

5. Conclusions

The surface limb of the Atlantic Meridional Overturning Circulation draws waters with negative PV from the southern hemisphere into the northern hemisphere. 32 Sv of this transport occurs in

the upper 600 m of the North Brazil Current (Johns et al. 1998; Schott et al. 1993). The negative PV of the southern hemisphere waters must be modified to enable crossing of the equator by more than a few Rossby deformation radii, and to ensure stability to symmetric instability. The details of the modification mechanism have been hitherto unknown; however, this work finds compelling theoretical and model based evidence that the excitement of symmetric instability is able to set the PV of cross-equatorial western boundary flows to zero.

Linear stability theory predicts that western boundary currents similar to the North Brazil Current should become symmetrically unstable upon crossing the equator. The theory is able to predict a growth rate, characteristic vertical length scale, and structure of the overturning cells which develop during the excitement of symmetric instability. The vertical length scale, however, is dependent upon the vertical eddy viscosity. Interpreting this result for real fluids presents a challenge; however, based on the work of Griffiths (2003a) it seems likely that secondary instabilities will play an important role. The equations of motion that govern the overturning streamfunction allow the overturning cells to extend into regions of positive PV. This allows the mixing of waters with positive and negative PV. Linear stability theory also predicts that a western boundary current such as the North Brazil Current may become symmetrically unstable at low latitudes south of the equator. Such instabilities are not considered in this work as the anomalous PV has different origins to that found in a cross-equatorial flow.

To verify the results of the linear stability analysis, and evaluate its applicability to more realistic currents, a simplified two-dimensional, meridionally symmetric, numerical model was developed. This model incorporates the effects of non-uniform stratification and vertical shear in the flow, which the linear stability analysis neglects. The model shows that the predictions of the linear stability analysis, including the development of stacked overturning cells with a characteristic vertical length scale, remain despite the added complexity. Surprisingly, there is no change in the evolution of

symmetric instability when the non-traditional component of the Coriolis force is included (e.g. Zeitlin 2018). By returning to linear stability theory we can see that this is not unexpected — the overturning streamfunction which drives the redistribution of PV, and the PV itself, are independent of the non-traditional Coriolis parameter. By introducing the complete Coriolis force, we introduce a natural meridional direction. This breaks the equivalence between meridional and zonal flows that exists on a traditional f-plane, meaning results derived for zonal flows will not necessarily hold true for meridional ones. More generally it was found that subjecting a flow to any irrotational force in the plane perpendicular to it, will not alter the evolution of symmetric instability within the flow.

The two-dimensional model neglects the β-effect and constrains the horizontal motions of the fluid flow. A further degree of realism was added by developing an idealized three-dimensional model of the tropical Atlantic which incorporates these effects. At the northern and southern boundaries of the model an idealized western boundary flow is prescribed. As the flow evolves two types of instability develop. The first is the spinning up of anticyclonic eddies as the fluid crosses the equator. This is a barotropic instability, and has been both observed and investigated in works by Edwards and Pedlosky (1998a,b) and Goes et al. (2009). The second instability is identified as symmetric instability for the following reasons:

- The characteristic timescale of the instability is of the same order of magnitude as that predicted by the linear stability analysis.

- The vertical length scale of the overturning cells that develop as a result of the instability agree with those predicted by the linear stability analysis.

- The instability is confined to regions of negative PV with minimal penetration of the instability into regions of positive PV. Eddies with strongly negative PV fuel its excitement.

- Waters leaving the region of instability have PV that is approximately neutrally stable to symmetric instability.

- The position of the instability is well explained by arguments about the latitudinal dependence of the growth rate of symmetric instability.

- The instability can be supressed in models with the imposition of a sufficiently high vertical viscosity, as predicted by linear stability theory.

- The power spectrum of the vertical kinetic energy in the standard three-dimensional integration, is similar to that in the corresponding two-dimensional integration. The latter unambiguously exhibits symmetric instability, due to the absence of barotropic eddies. Moreover, the viscous two-dimensional integration, which shows no signs of symmetric instability, does not show the same power spectrum flattening seen in the two-dimensional and three-dimensional integrations which are identified as exhibiting symmetric instability.

Yankovsky and Legg (2019) suggest that analyzing the power spectrum of vertical kinetic energy may be a useful technique in identifying signs of symmetric instability in more complex and noisy models — a suggestion that our work strongly supports.

Our initial hypothesis was that the excitement of symmetric instability could lead to transformation of fluid between different water mass classes, leading to changes in our understanding of the overturning budget of the AMOC. However, an analysis of the water mass formation rates in the surface, pycnocline and deep ocean suggests that the net contribution of symmetric instability to the budget is likely to be small. Over short timescales, symmetric instability appears to drive transformations between these density classes of around ± 2 Sv, but as high as ± 4 Sv at times; however, over longer timescales these transformations seem to average out to a net rate that is generally around ± 1 Sv in the surface and pycnocline waters and ± 2 Sv in the deep waters. A limitation of the water

mass transformation calculations is the viscosity dependence of the overturning cells generated by the instability. In the ocean, we would expect the overturning cells to have a smaller vertical extent of around 15 m, which would likely alter the formation rates seen here. Furthermore, Yankovsky and Legg (2019) find that the mixing of waters by symmetric instability is largely adiabatic and that secondary shear instabilities are found to cause irreversible mixing. This suggests that using a high resolution model is key to being able to reliably measure water mass transformations. Achieving a high enough resolution with the fully three-dimensional model may be difficult; however, the two-dimensional models could prove useful for attaining better estimates. The low resolution of the models used here may also lead to incomplete representation of the symmetric instability if the overturning cells are oriented with isopycnals. However, this is unlikely to change the key result that symmetric instability is efficient at removing anomalous PV originating from opposing hemispheres.

The work presented here has focused on currents dynamically similar to the North Brazil Current. It is not clear what the role of symmetric instability in the modification of the- PV of other cross-equatorial currents is. For instance, the Somali current is an intense, seasonal, wind-driven, cross-equatorial western boundary current in the Indian ocean. In theory, symmetric instability could also be at play there, as long as the relative vorticity is large enough to produce a significant growth rate. The return flow of the AMOC takes the form of a deep western boundary current, with maximum speeds around a fifth of what is observed in the upper ocean. We have performed preliminary investigations into symmetric instability in an idealized deep western boundary current, with initial results suggesting it may too be susceptible to symmetric instability. The instability however is much less explosive, with a characteristic timescale of several months. It is not yet clear whether the instability is sufficient to dominate over boundary layer processes in the modification of anomalous PV.

This work has not explored the dynamics of waters with positive PV in the southern hemisphere. Against the western boundary is a region of positive absolute vorticity. However, it is likely that dissipative fluxes of PV dominate the PV balance in the region, potentially providing a more efficient mechanism by which anomalous PV can be removed (Edwards and Pedlosky 1998a).

Acknowledgments. We are grateful for the financial support of the Natural Environment Research Council NE/L002612/1. This work made use of the ARCHER UK National Supercomputing Service (`http://www.archer.ac.uk`). We would also like to thank Elizabeth Yankovsky and an anonymous reviewer, for their detailed comments which improved the manuscript, and Brian Hoskins for discussions on the different taxonomic systems used to describe symmetric instabilities. This work used data that was collected and made freely available by the International Argo Program and the national programs that contribute to it. (`https://argo.ucsd.edu`, `https://www.ocean-ops.org`). The Argo Program is part of the Global Ocean Observing System.

Data availability statement. The jupyter notebook used in the linear stability analysis, and the MITgcm configuration files used for the two-dimensional and three-dimensional model integrations are available from `https://github.com/fraserwg/SIInCrossEquatorialWBCs`, and are archived on Zenodo (Goldsworth et al. 2021). Animated versions of snapshot figures are available in the supplementary materials.

References

AMS, 2020: AMS glossary of meteorology: symmetric instability. American Meteorological Society, URL https://glossary.ametsoc.org/w/index.php?title=Symmetric_instability&oldid=8723.

Argo, 2000: Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). SEANOE, URL https://doi.org/10.17882/42182.

Bachman, S. D., and J. R. Taylor, 2014: Modelling of partially-resolved oceanic symmetric instability. *Ocean Modelling*, **82**, 15–27, doi:10.1016/j.ocemod.2014.07.006, URL http://dx.doi.org/10.1016/j.ocemod.2014.07.006.

Bower, A., and Coauthors, 2019: Lagrangian Views of the Pathways of the Atlantic Meridional Overturning Circulation. *Journal of Geophysical Research: Oceans*, **124 (8)**, 5313–5335, doi:10.1029/2019JC015014.

Brannigan, L., 2016: Intense submesoscale upwelling in anticyclonic eddies. *Geophysical Research Letters*, **43 (7)**, 3360–3369, doi:10.1002/2016GL067926.

Bryden, H. L., W. E. Johns, B. A. King, G. McCarthy, E. L. McDonagh, B. I. Moat, and D. A. Smeed, 2020: Reduction in ocean heat transport at 26°N since 2008 cools the eastern subpolar gyre of the North Atlantic Ocean. *Journal of Climate*, **33 (5)**, 1677–1689, doi:10.1175/JCLI-D-19-0323.1.

Buckley, M. W., and J. Marshall, 2015: Observations, inferences, andmechanisms of the Atlantic Meridional Overturning Circulation: A review. *Reviews of Geophysics*, **54**, 5–63, doi:10.1002/2015RG000493.Received.

Castelão, G. P., and W. E. Johns, 2011: Sea surface structure of North Brazil Current rings derived from shipboard and moored acoustic Doppler current profiler observations. *Journal of Geophysical Research: Oceans*, **116 (1)**, doi:10.1029/2010JC006575.

Eden, G., and J. Willebrand, 1999: Neutral density revisited. *Deep-Sea Research Part II: Topical Studies in Oceanography*, **46 (1-2)**, 33–54, doi:10.1016/S0967-0645(98)00113-1.

Edwards, C. A., and J. Pedlosky, 1998a: Dynamics of nonlinear cross-equatorial flow. Part I: Potential vorticity transformation. *Journal of Physical Oceanography*, **28 (12)**, 2382–2406, doi:10.1175/1520-0485(1998)028<2382:DONCEF>2.0.CO;2.

838 Edwards, C. A., and J. Pedlosky, 1998b: Dynamics of Nonlinear Cross-Equatorial Flow. Part
839 II: The Tropically Enhanced Instability of the Western Boundary Current. *Journal of Physical*
840 *Oceanography*, **28 (12)**, 2407–2417, doi:10.1175/1520-0485(1998)028<2407:DONCEF>2.0.
841 CO;2.

842 Fonseca, C. A., G. J. Goni, W. E. Johns, and E. J. Campos, 2004: Investigation of the North Brazil
843 Current retroflection and North Equatorial Countercurrent variability. *Geophysical Research*
844 *Letters*, **31 (21)**, 1–5, doi:10.1029/2004GL020054.

845 Goes, M. P., D. P. Marshall, and I. Wainer, 2009: Eddy Formation in the Tropical Atlantic
846 Induced by Abrupt Changes in the Meridional Overturning Circulation. *Journal of Physical*
847 *Oceanography*, **39 (11)**, 3021–3031, doi:10.1175/2009JPO4004.1.

848 Goldsworth, F. W., D. P. Marshall, and H. L. Johnson, 2021: GitHub repository for "Symmetric
849 instability in cross equatorial western boundary currents". Zenodo, URL https://doi.org/10.5281/
850 zenodo.4650332, doi:10.5281/zenodo.4650332.

851 Griffies, S. M., and R. W. Hallberg, 2000: Biharmonic friction with a Smagorinsky-like viscosity
852 for use in large-scale eddy-permitting ocean models. *Monthly Weather Review*, **128 (8)**, 2935–
853 2946, doi:10.1175/1520-0493(2000)128<2935:bfwasl>2.0.co;2.

854 Griffiths, S. D., 2003a: Nonlinear vertical scale selection in equatorial inertial instability. *Journal of*
855 *the Atmospheric Sciences*, **60 (7)**, 977–990, doi:10.1175/1520-0469(2003)060<0977:NVSSIE>
856 2.0.CO;2.

857 Griffiths, S. D., 2003b: The nonlinear evolution of zonally symmetric equatorial inertial instability.
858 *Journal of Fluid Mechanics*, **474 (474)**, 245–273, doi:10.1017/S0022112002002586.

Haine, T. W. N., and J. Marshall, 1998: Gravitational, Symmetric, and Baroclinic Instability of the Ocean Mixed Layer. *Journal of Physical Oceanography*, **28 (4)**, 634–658, doi:10.1175/1520-0485(1998)028<0634:gsabio>2.0.co;2.

Holton, J. R., and G. J. Hakim, 2013: Mesoscale Circulations. *An Introduction to Dynamic Meteorology*, Elsevier, 279–323, doi:10.1016/B978-0-12-384866-6.00009-X, URL https://linkinghub.elsevier.com/retrieve/pii/B978012384866600009X.

Hoskins, B. J., 1974: The role of potential vorticity in symmetric stability and instability. *Quarterly Journal of the Royal Meteorological Society*, **100 (425)**, 480–482, doi:10.1002/qj.49710042520.

Hua, B. L., D. W. Moore, and S. Le Gentil, 1997: Inertial nonlinear equilibration of equatorial flows. *Journal of Fluid Mechanics*, **331**, 345–371, doi:10.1017/S0022112096004016.

Jackson, L. C., R. Kahana, T. Graham, M. A. Ringer, T. Woollings, J. V. Mecking, and R. A. Wood, 2015: Global and European climate impacts of a slowdown of the AMOC in a high resolution GCM. *Climate Dynamics*, **45 (11-12)**, 3299–3316, doi:10.1007/s00382-015-2540-2, URL http://dx.doi.org/10.1007/s00382-015-2540-2.

Jochum, M., and P. Malanotte-Rizzoli, 2003: On the generation of North Brazil Current rings. *Journal of Marine Research*, **61 (2)**, 147–173, doi:10.1357/002224003322005050.

Johns, W. E., T. N. Lee, R. C. Beardsley, J. Candela, R. Limeburner, and B. Castro, 1998: Annual cycle and variability of the North Brazil current. *Journal of Physical Oceanography*, **28 (1)**, 103–128, doi:10.1175/1520-0485(1998)028<0103:ACAVOT>2.0.CO;2.

Joshi, M., 1994: Orographic infliuences in the atmosphere of Mars. Ph.D. thesis, University of Oxford.

Killworth, P. D., 1991: Cross-equatorial Geostrophic Adjustment. *Journal of Physical Oceanography*, **21 (10)**, 1581–1601, doi:10.1175/1520-0485(1991)021<1581:CEGA>2.0.CO;2.

Kloosterziel, R. C., G. F. Carnevale, and P. Orlandi, 2017: Equatorial inertial instability with full Coriolis force. *Journal of Fluid Mechanics*, **825**, 69–108, doi:10.1017/jfm.2017.377.

Marshall, D. P., and H. R. Pillar, 2011: Momentum balance of the Wind-Driven and Meridional Overturning Circulation. *J.Phys.Oceanogr.*, **41**, 960–978, doi:10.1175/201.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997: A finite-volume, incompressible navier stokes model for, studies of the ocean on parallel computers. *Journal of Geophysical Research C: Oceans*, **102 (C3)**, 5753–5766, doi:10.1029/96JC02775.

Plougonven, R., and V. Zeitlin, 2009: Nonlinear development of inertial instability in a barotropic shear. *Physics of Fluids*, **21 (10)**, doi:10.1063/1.3242283.

Prather, M. J., 1986: Numerical advection by conservation of second-order moments. *Journal of Geophysical Research*, **91 (D6)**, 6671, doi:10.1029/JD091iD06p06671, URL http://doi.wiley.com/10.1029/JD091iD06p06671.

Ribstein, B., V. Zeitlin, and A. S. Tissier, 2014: Barotropic, baroclinic, and inertial instabilities of the easterly Gaussian jet on the equatorial β-plane in rotating shallow water model. *Physics of Fluids*, **26 (5)**, doi:10.1063/1.4875030.

Rodwell, M. J., and B. J. Hoskins, 1995: A model of the asian summer monsoon.part ii: Cross-equatorial flow and pv behavior. *Journal of Atmospheric Sciences*, **52 (9)**, 1341 – 1356, doi:10.1175/1520-0469(1995)052<1341:AMOTAS>2.0.CO;2, URL https://journals.ametsoc.org/view/journals/atsc/52/9/1520-0469_1995_052_1341_amotas_2_0_co_2.xml.

Schott, F., J. Fischer, J. Reppin, and U. Send, 1993: On mean and seasonal currents and transports at the western boundary of the equatorial Atlantic. *Journal of Geophysical Research*, **98 (C8)**, doi:10.1029/93jc01287.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations. *Monthly Weather Review*, **91 (3)**, 99–164, doi:10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Srokosz, M., M. Baringer, H. Bryden, S. Cunningham, T. Delworth, S. Lozier, J. Marotzke, and R. Sutton, 2012: Past, present, and future changes in the atlantic meridional overturning circulation. *Bulletin of the American Meteorological Society*, **93 (11)**, 1663–1676, doi:10.1175/BAMS-D-11-00151.1.

Stewart, A. L., and P. J. Dellar, 2011: The rôle of the complete Coriolis force in cross-equatorial flow of abyssal ocean currents. *Ocean Modelling*, **38 (3-4)**, 187–202, doi:10.1016/j.ocemod.2011.03.001.

Stone, P. H., 1966: On Non-Geostrophic Baroclinic Stability. *Journal of the Atmospheric Sciences*, **23 (4)**, 390–400, doi:10.1175/1520-0469(1966)023<0390:ONGBS>2.0.CO;2, URL http://journals.ametsoc.org/doi/10.1175/1520-0469(1966)023{\%}3C0390:ONGBS{\%}3E2.0.CO;2.

Talley, L. D., G. L. Pickard, W. J. Emery, and J. H. Swift, 2011: Atlantic Ocean. *Descriptive Physical Oceanography*, Elsevier, chap. 9, 245–301, doi:10.1016/B978-0-7506-4552-2.10009-5, URL https://linkinghub.elsevier.com/retrieve/pii/B9780750645522100095.

Taylor, J. R., and R. Ferrari, 2009: On the equilibration of a symmetrically unstable front via a secondary shear instability. *Journal of Fluid Mechanics*, **622**, 103–113, doi:10.1017/

S0022112008005272.

Thomas, L. N., J. R. Taylor, R. Ferrari, and T. M. Joyce, 2013: Symmetric instability in the Gulf Stream. *Deep-Sea Research Part II: Topical Studies in Oceanography*, **91**, 96–110, doi:10.1016/j.dsr2.2013.02.025, URL https://www.sciencedirect.com/science/article/pii/S0967064513000829.

Vianna, M., and V. de Menezes, 2003: A seasonal and interannual study of the western equatorial atlantic upper thermocline circulation variability. *Interhemispheric Water Exchange in the Atlantic Ocean*, G. Goni, and P. Malanotte-Rizzoli, Eds., Elsevier Oceanography Series, Vol. 68, Elsevier, 137 – 173, doi:https://doi.org/10.1016/S0422-9894(03)80145-9, URL http://www.sciencedirect.com/science/article/pii/S0422989403801459.

Walin, G., 1982: On the relation between sea-surface heat flow and thermal circulation in the ocean. *Tellus*, **34 (2)**, 187–195, doi:10.3402/tellusa.v34i2.10801.

Williams, R. G., and M. J. Follows, 2011: *Ocean Dynamics and the Carbon Cycle*. Cambridge University Press, Cambridge.

Yankovsky, E., and S. Legg, 2019: Symmetric and baroclinic instability in dense shelf overflows. *Journal of Physical Oceanography*, **49 (1)**, 39–61, doi:10.1175/JPO-D-18-0072.1.

Zeitlin, V., 2018: Symmetric instability drastically changes upon inclusion of the full Coriolis force. *Physics of Fluids*, **30 (6)**, 061 701–1–4, doi:10.1063/1.5031099.

LIST OF FIGURES

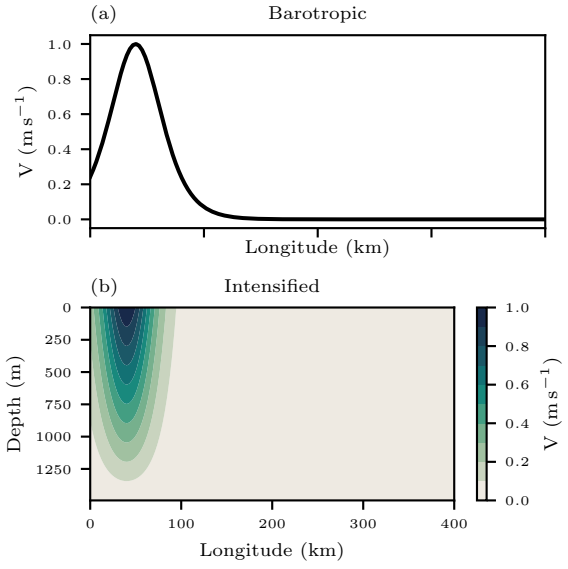
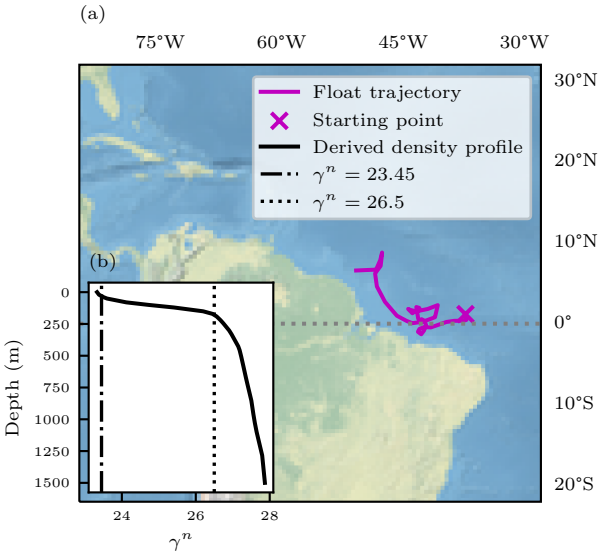


FIG. 1. Velocity structure of the (a) barotropic and (b) surface intensified Bickley jet, with $V_0 = 1$ m s^{-1}, $x_{mid} = 40$ km, $\delta_b = 30$ km and $H = 1,500$ m.



1004 FIG. 2. The map (a) shows the trajectory of the Argo float on which the initial stratification of the model
1005 is based. Inset (b) is the neutral density profile used to initialise the model, based on observations from the
1006 Argo float. The vertical dash-dotted and dotted lines in (b) mark the γ^n boundaries used in the water mass
1007 transformation calculations discussed in section 4b.

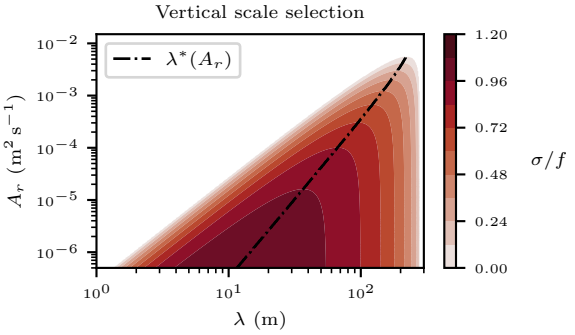


FIG. 3. The growth rate of symmetric instability in a barotropic Bickley jet is plotted as a function of the vertical wavelength of the instability and the viscosity of the fluid. Overlain (dash-dotted line) is the surface along which the growth rate is maximised for a given viscosity. Regions in white correspond to exponentially decaying modes. There are no unstable modes at wavelengths greater than 285 m at all vertical viscosities.

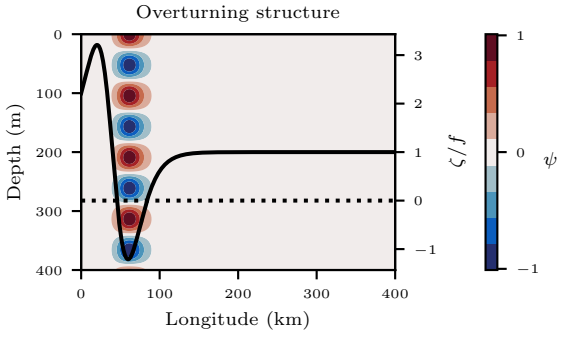
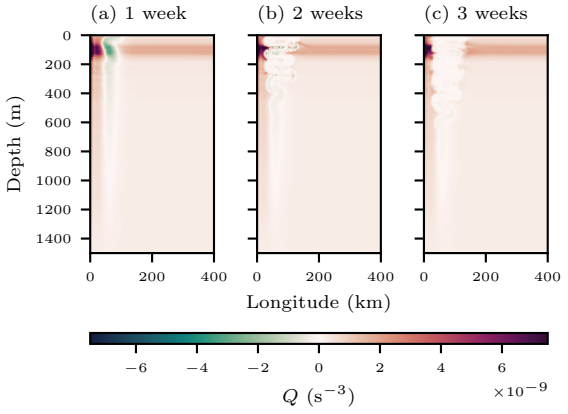


FIG. 4. Contours show the zonal overturning streamfunction generated by symmetric instability in a barotropic Bickley jet, as predicted by the linear stability analysis. Overlain (the solid black line) is the absolute vorticity of the initial jet. The absolute vorticity is equal to zero along the dotted line. The depth coordinate may be shifted by an arbitrary amount due to the assumption of normal mode solutions in the vertical.



1016 FIG. 5. Snapshots of PV over time for the two-dimensional standard (no-slip) run. PV is shown as a function of
1017 depth and longitude. An animated version of this figure is available in the supplementary materials as figure 5S.

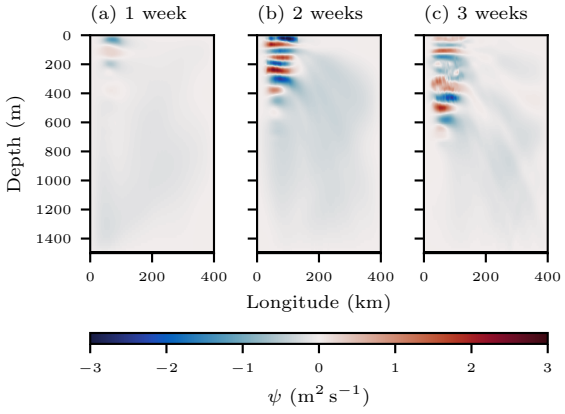


FIG. 6. Snapshots of the zonal overturning streamfunction generated by symmetric instability in the two-dimensional standard (no-slip) run. The 5 day moving average of the streamfunction has been taken to mask the effects of an inertial wave.

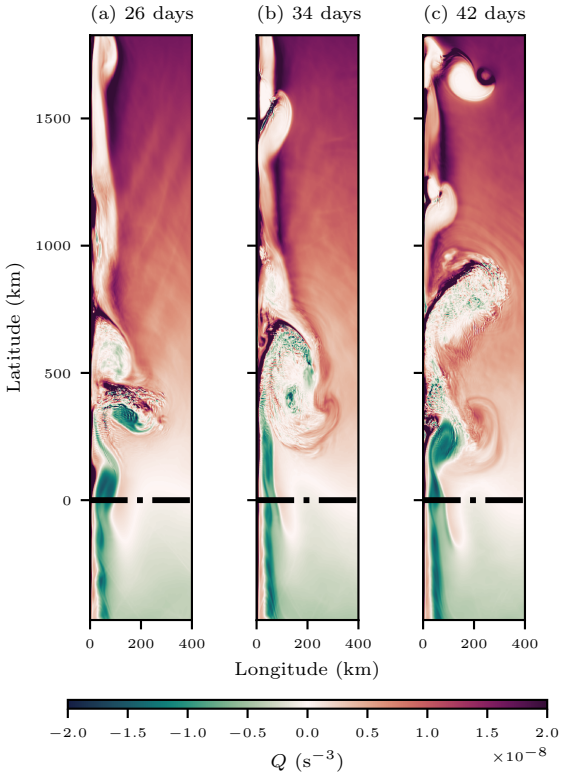
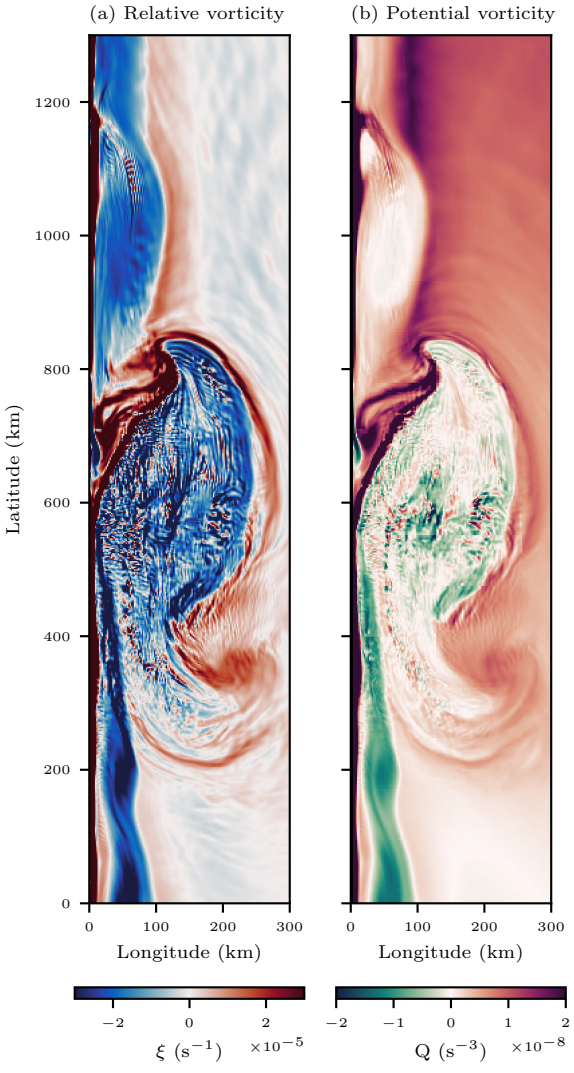


FIG. 7. Snapshots of PV at 50 m depth for the standard (no-slip) three-dimensional model. Note that this image, and all other snapshots at 50 m depth, do not show the full model domain (the sponge regions and the area East of 400 km have been cropped out.) An animated version of this figure is available in the supplementary materials as figure 7S.



1025 FIG. 8. Relative and potential vorticity at 50 m depth in the same region north of the equator, using fields from
1026 the standard (no-slip) three-dimensional models.

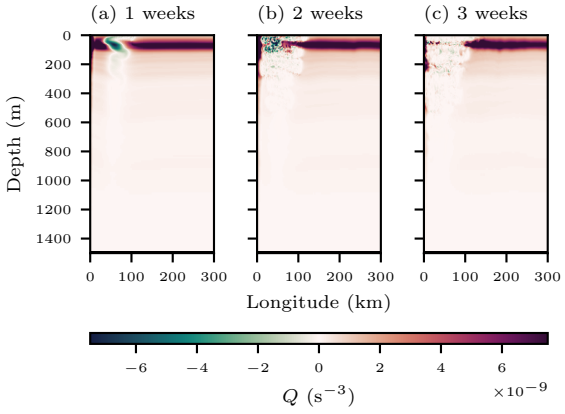


FIG. 9. Snapshots of PV as a function of depth and longitude at 750 km north of the equator, in the standard (no-slip) three-dimensional model. An animated version of this figure is available in the supplementary materials as figure 9S.

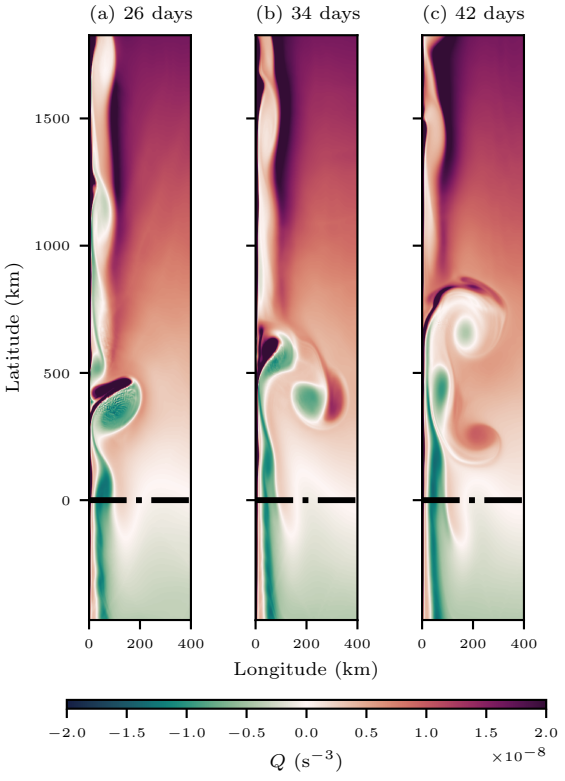


FIG. 10. Snapshots of PV at 50 m depth for the no-slip viscous three-dimensional model. An animated version of this figure is available in the supplementary materials as figure 10S.

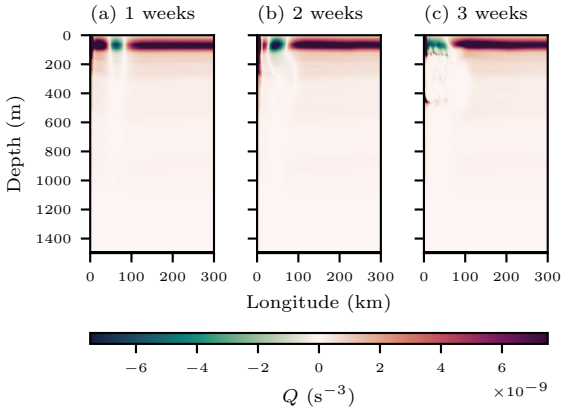


FIG. 11. Snapshots of PV as a function of depth and longitude at 750 km north of the equator, in the no-slip viscous three-dimensional model. An animated version of this figure is available in the supplementary materials as figure 11S.

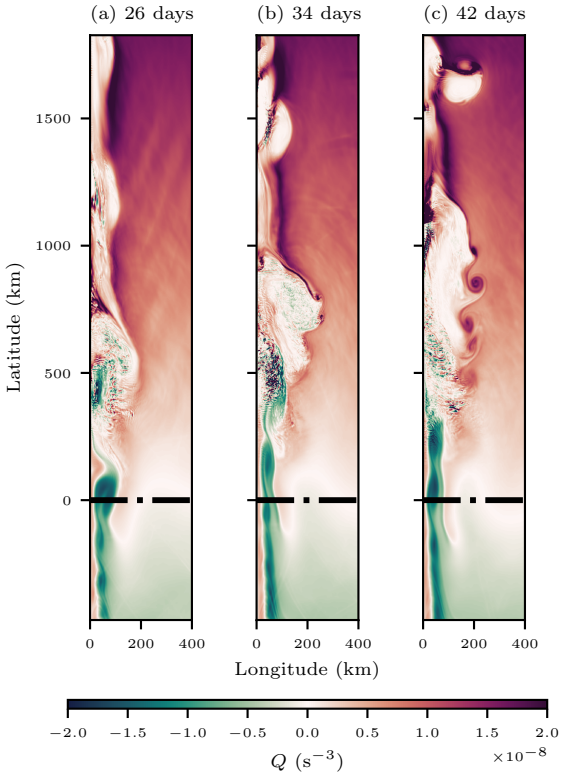


FIG. 12. Snapshots of PV at 50 m depth for the free-slip three-dimensional model. An animated version of this figure is available in the supplementary materials as figure 12S.

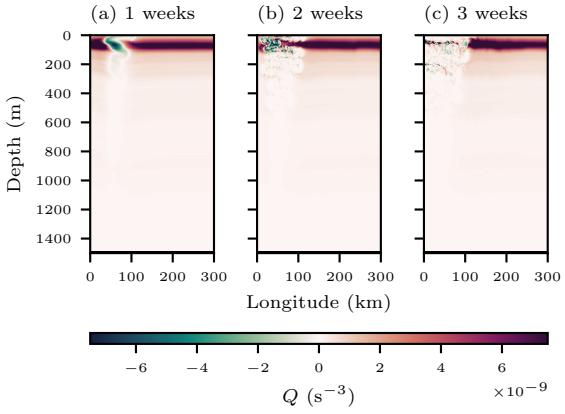


FIG. 13. Snapshots of PV as a function of depth and longitude at 750 km north of the equator, in the free-slip three-dimensional model. An animated version of this figure is available in the supplementary materials as figure 13S.

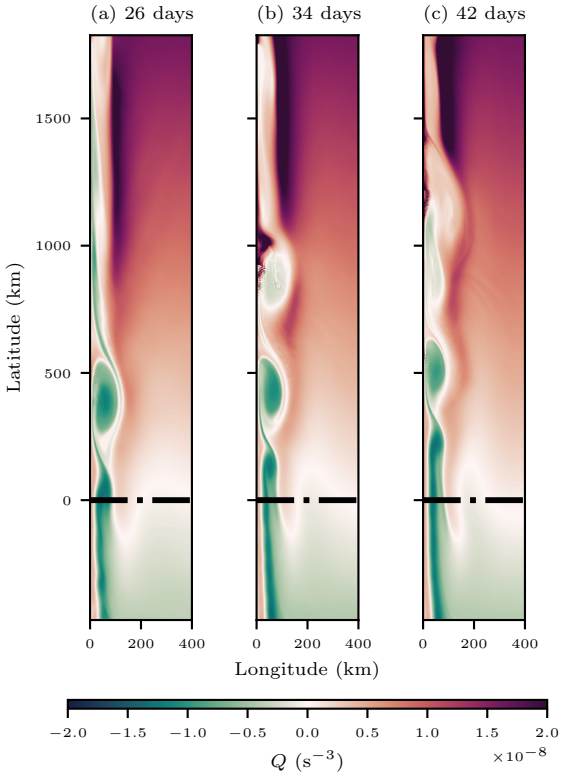


FIG. 14. Snapshots of PV at 50 m depth for the free-slip viscous three-dimensional model. An animated version of this figure is available in the supplementary materials as figure 14S.

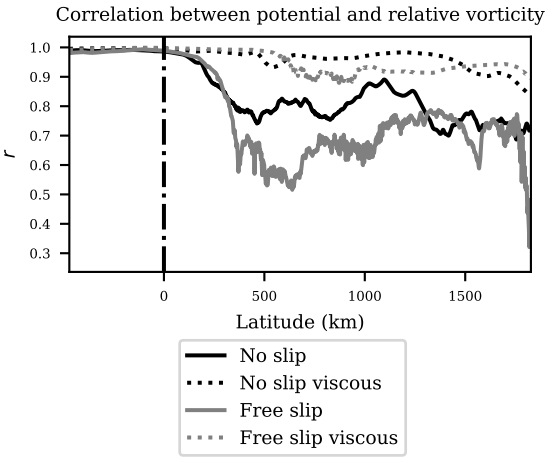


FIG. 15. Correlations between relative and potential vorticity as a function of latitude. Shown for integrations with a standard (solid lines) or enhanced (dotted lines) vertical viscosity, and for both no-slip (black lines) and free-slip (grey lines) models.

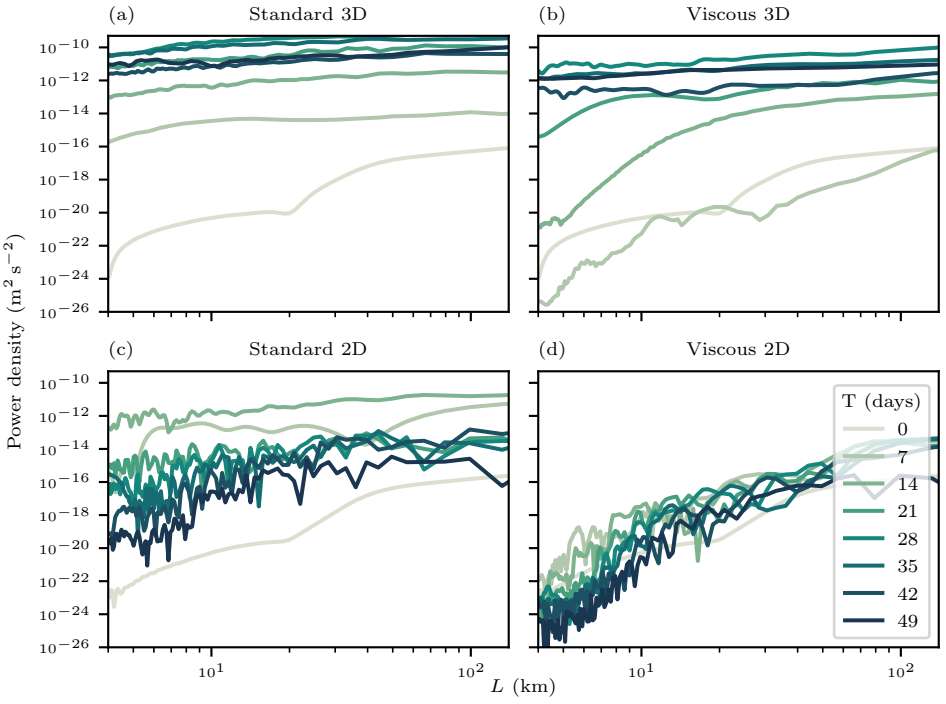


FIG. 16. Spectral distribution of vertical kinetic energy for the (a) standard three-dimensional, (b) viscous three-dimensional, (c) standard two-dimensional, and (d) viscous two dimensional models. Note that the vertical viscosity used in the viscous two-dimensional integration is 6×10^{-3} m^2 s^{-1} compared to 10^{-2} m^2 s^{-1} for its three-dimensional counter part. Darkening line colour corresponds to later model times.

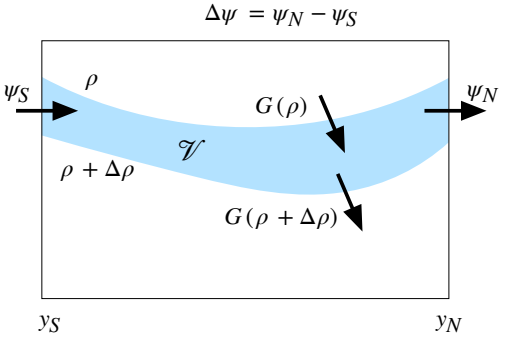
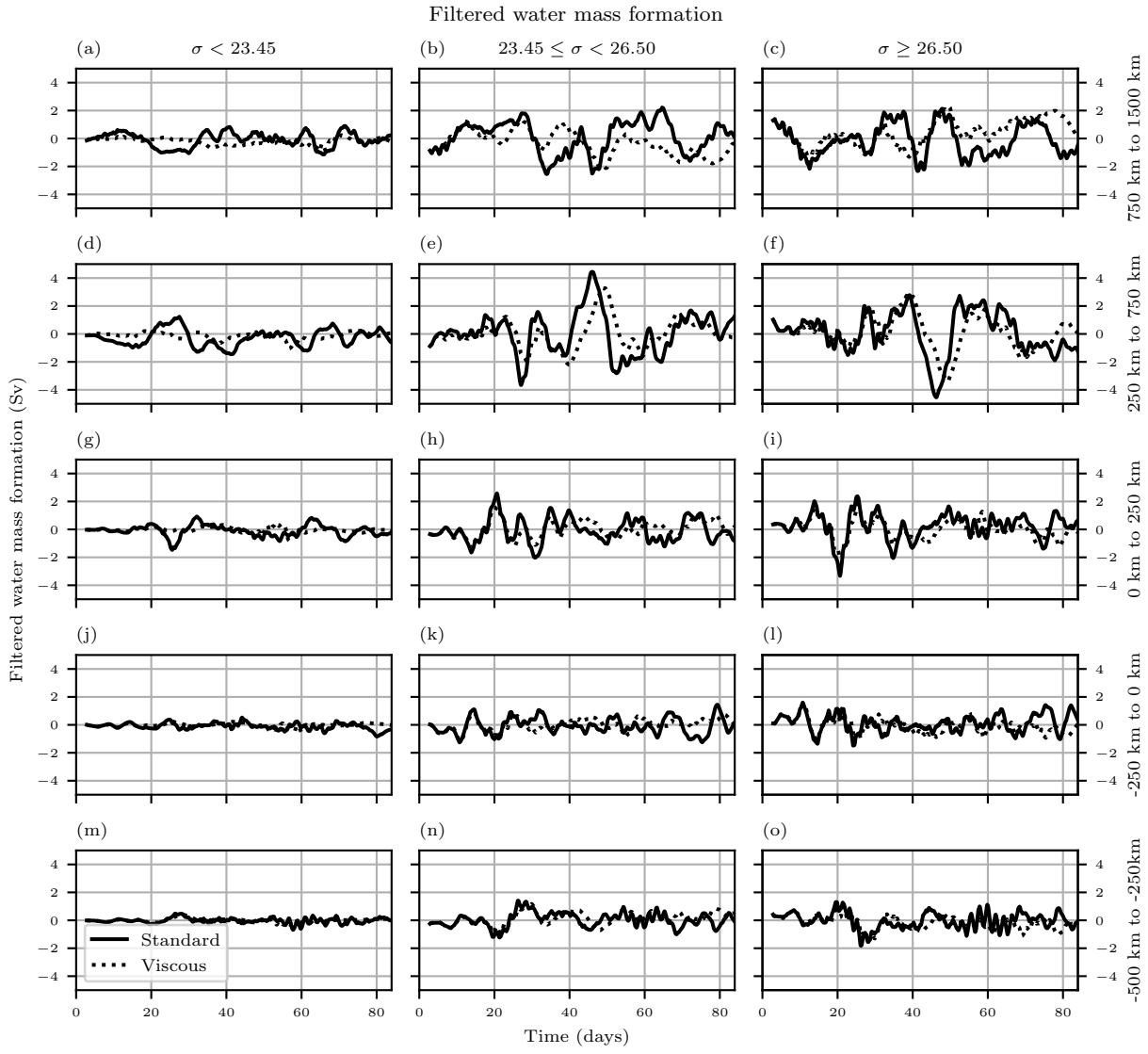
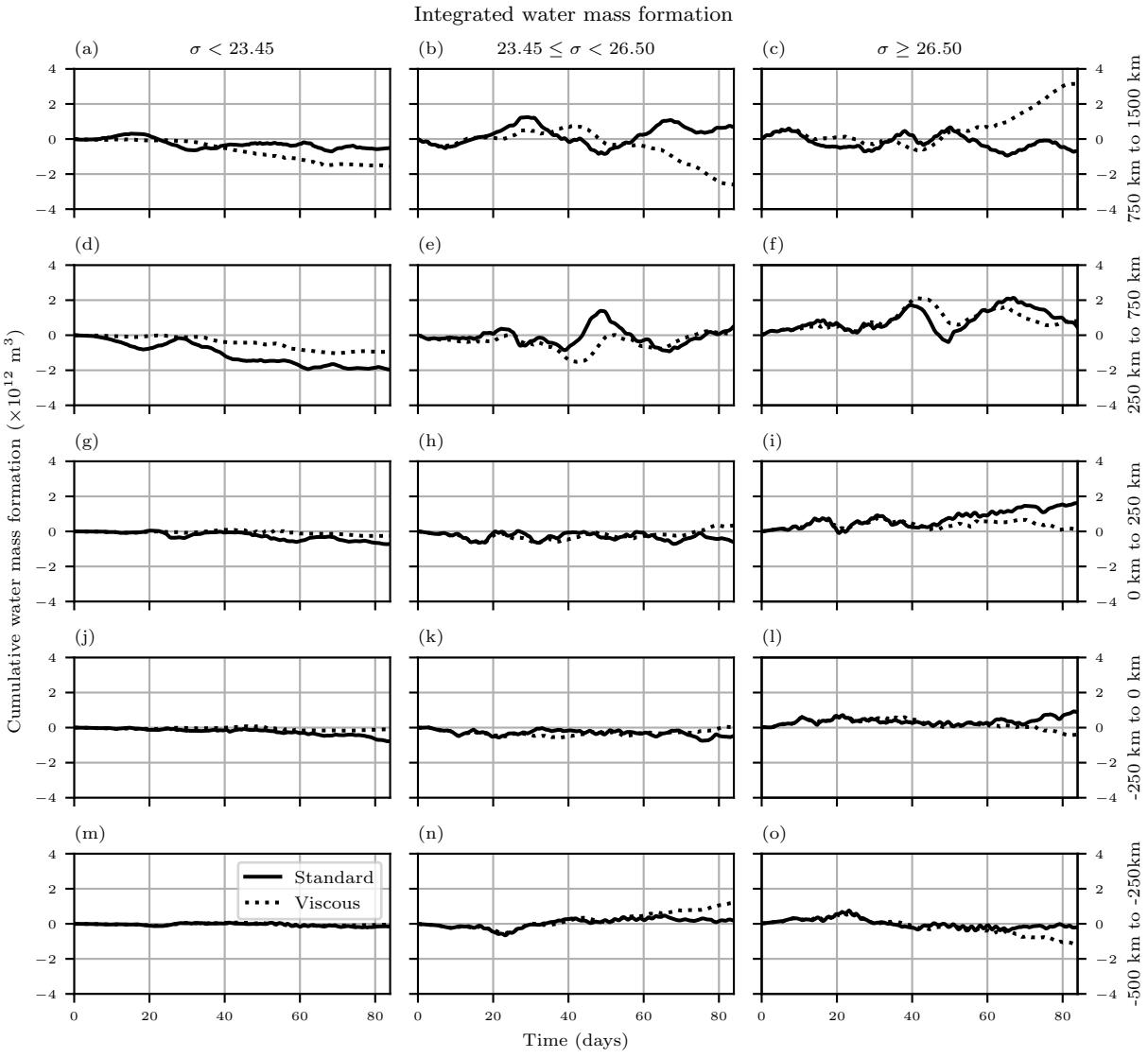


FIG. 17. Sketch of the processes contributing to water mass formation in the Walin framework. Adapted from Williams and Follows (2011).



1051 FIG. 18. Water mass formation rates for the surface, pycnocline and deep ocean (columns) within different
1052 latitude bands (rows). Rates are shown for the three-dimensional standard (solid lines) and viscous (dotted lines)
1053 no-slip models. High frequency variability was filtered from the rates by taking a 2.9 day rolling average.



1054 FIG. 19. Cumulatively integrated water mass formation rates for the surface, pycnocline and deep ocean
1055 (columns) within different latitude bands (rows). The net formation is plotted for the three-dimensional standard
1056 (solid lines) and viscous (dotted lines) no-slip models.