JGR Oceans

RESEARCH ARTICLE
10.1029/2023JC020795

Surface Heat Fluxes Drive a Two-Phase Response in Southern Ocean Mode Water Stratification

Ciara Pimm[1], Richard G. Williams[1], Dani Jones[2,3], and Andrew J. S. Meijers[2]

[1]Department of Earth, Ocean and Ecological Sciences, School of Environmental Sciences, University of Liverpool, Liverpool, UK, [2]British Antarctic Survey, NERC, UKRI, Cambridge, UK, [3]Cooperative Institute for Great Lakes Research, University of Michigan, Ann Arbor, MI, USA

Abstract Subantarctic mode waters have low stratification and are formed through subduction from thick winter mixed layers in the Southern Ocean. To investigate how surface forcing affects the stratification in mode water formation regions in the Southern Ocean, a set of adjoint sensitivity experiments are conducted. The objective function is the annual-average stratification over the mode water formation region, which is evaluated from potential temperature and salinity adjoint sensitivity experiments. The analysis of impacts, from the product of sensitivities and forcing variability, identifies the separate effects of the wind stress, heat flux, and freshwater flux, revealing that the dominant control on stratification is from surface heat fluxes, as well as a smaller effect from zonal wind stress. The adjoint sensitivities of stratification to surface heat flux reveal a surprising change in sign over 2 years lead time: surface cooling leads to the expected initial local decrease in stratification, but there is a delayed response leading to an increase in stratification. This delayed response in stratification involves effective atmospheric damping of the surface thermal contribution, so that eventually the oppositely-signed advective haline contribution dominates. This two-phase response of stratification is found to hold over mode water formation regions in the South Indian and Southeast Pacific sectors of the Southern Ocean, where there are strong advective flows linked to the Antarctic Circumpolar Current.

Plain Language Summary The Southern Ocean, surrounding the Antarctic continent, plays an important role in the uptake and transport of heat and carbon. Subantarctic mode waters, which are characterized by their low stratification, play an important role in this uptake of heat and carbon, and therefore the factors impacting their properties need to be properly understood. To understand how surface forcing affects Subantarctic mode waters, sensitivity studies are conducted in an ocean state estimate, which consider the relative importance of surface heat flux, freshwater flux, and wind stresses on the stratification of mode waters. Surface heat flux has the largest impact on mode water formation both on seasonal and longer interannual timescales. Initially, surface heat loss leads to a decrease in stratification in the mode waters. However, there is a delayed response where the surface temperature response is effectively damped by the atmosphere and there is an opposing-signed salinity response advected into the region, leading to a subsequent increase in stratification in the mode waters.

Key Points:

- The sensitivity of Southern Ocean mode water stratification to surface heat flux changes sign over time
- Surface heat loss leads to an initial decrease in stratification in the mode waters
- Surface heat loss leads to a delayed restratification due to a haline contribution after a thermal contribution is effectively damped

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

C. Pimm,
c.pimm@liverpool.ac.uk

Citation:

Pimm, C., Williams, R. G., Jones, D., & Meijers, A. J. S. (2024). Surface heat fluxes drive a two-phase response in Southern Ocean mode water stratification. *Journal of Geophysical Research: Oceans*, *129*, e2023JC020795. https://doi.org/10.1029/2023JC020795

Received 6 DEC 2023
Accepted 29 FEB 2024

Author Contributions:

Conceptualization: Ciara Pimm
Data curation: Ciara Pimm
Formal analysis: Ciara Pimm
Investigation: Ciara Pimm, Richard G. Williams
Methodology: Ciara Pimm, Richard G. Williams
Supervision: Richard G. Williams, Dani Jones, Andrew J. S. Meijers
Writing – original draft: Ciara Pimm
Writing – review & editing: Ciara Pimm, Richard G. Williams, Dani Jones, Andrew J. S. Meijers

1. Introduction

The Southern Ocean is globally important due to its role in the uptake and transport of heat and anthropogenic carbon (Frölicher et al., 2015; Williams et al., 2023). The Southern Ocean connects the three main ocean basins, Indian, Pacific, and Atlantic through the Antarctic Circumpolar Current (ACC). The Southern Ocean overturning circulation plays a pivotal role in the global overturning circulation (Marshall & Speer, 2012), involving the time-mean circulation and time-varying eddies providing vertical and meridional transport across the ACC. Ventilation involves the process of surface water moving from the mixed layer to the thermocline and deep oceans, as part of the overturning, allowing the transport of atmospheric properties into the ocean interior (Sallée et al., 2010; Williams & Meijers, 2019). Ventilation can occur in many ways in the Southern Ocean ranging from seasonal local subduction, subtropical gyre transport, and the zonally-averaged overturning circulation and mesoscale-eddy exchange along isopycnals (Morrison et al., 2022).

Subantarctic Mode Water (SAMW) is a subsurface water mass, found equatorward of the fronts of the ACC, characterized by its low potential vorticity (PV) (McCartney, 1977). SAMW is important in the uptake and transport of heat (Armour et al., 2016; Gao et al., 2018; Haumann et al., 2016), carbon (Gruber et al., 2019), and

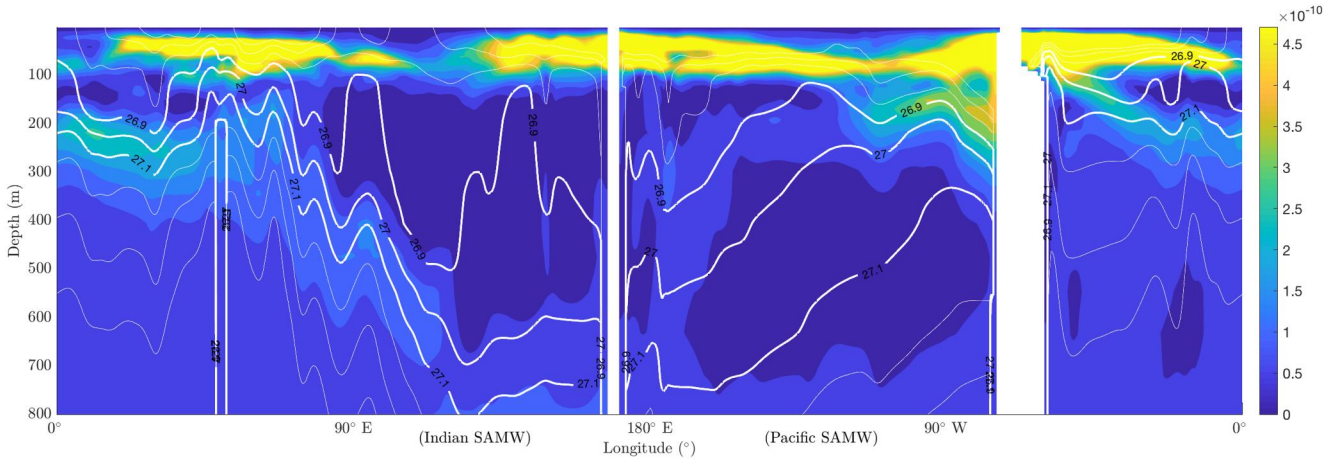

Figure 1. Annual-average potential vorticity (PV) for year 2011 (color shade in $m^{-1}\ s^{-1}$) at 45° South with neutral density surface contours (kg m^{-3}, white), with thicker contours [26.9, 27.0, 27.1] bounding areas of low PV over longitude (°) and depth (m). Subantarctic mode water is contained within the low PV waters within 26.8–27.2 kg m^{-3} (Sallée et al., 2010), with the Indian and Pacific formation sites labeled.

nutrients (Hendry & Brzezinski, 2014). SAMW formation sites are identified by their anomalously deep winter mixed layers (Herraiz-Borreguero & Rintoul, 2011) and are formed during the seasonal cycle of the mixed layer through the subduction processes (Kwon et al., 2013). Over austral winter there is buoyancy loss from the surface ocean due to surface cooling. High buoyancy loss leads to convective overturning, which creates deep mixed layers in winter, often associated with the passage of storms (Ogle et al., 2018). The seasonal mixed layers shoal in spring and leave behind a pool of low PV water (Figure 1), which may be subducted into the permanent pycnocline. Pools of low PV water in different Southern Ocean basins occur over density surface ranges from 26.8 to 27.2 kg m^{-3} (Figure 1), with lighter water masses in the South Indian and denser in the South Pacific (Sallée et al., 2010). Recent property variability has been observed in SAMW, which has freshened (Haumann et al., 2016) and thickened (Gao et al., 2018) over the late 1900s and early 2000s.

The main aim of this study is to address the physical processes that affect the stratification of mode waters in formation regions in the Southern Ocean and the role of different surface forcings and their effects with time. The main drivers of mode water formation are largely ascribed to the instantaneous forcing, but there is a secondary, largely unquantified, role for advection from upstream regions (Cerovečki & Meijers, 2021). There are particular uncertainties with the role of upstream forcing in setting SAMW properties.

To understand the detailed processes that control the properties of mode waters and their characteristic low stratification, adjoint sensitivity studies are completed. In Section 2, the experimental design is discussed, which describes the objective function of annual average large-scale PV, which is a measure of stratification, over the chosen South Indian control volume for the adjoint sensitivity experiments. In Section 3, linear adjoint sensitivities of large-scale PV to different surface forcings are considered and their relative magnitudes are compared using impacts. Section 4 discusses the time dependence of the adjoint sensitivities of large-scale PV to surface heat flux. Forward heat flux perturbation experiments are employed to clarify the physical understanding found from the adjoint sensitivities. The generality of our model analyses of how surface heat fluxes affect mode water stratification is considered by comparing both the South Indian and Southeast Pacific mode water formation regions.

2. Experimental Design

2.1. The ECCO State Estimate

Forward and adjoint modeling using the ECCO (Estimating the Circulation and Climate of the Ocean) state estimate (ECCO Consortium et al., 2020) are completed to explore the geographical and time-varying influences of surface forcing, particularly surface heat flux, on stratification within the South Indian mode water formation pool. ECCO is a physically-consistent ocean state estimate that conserves heat, salt, mass, and momentum and thus does not have arbitrary sources or sinks (Forget, Campin, et al., 2015). An advantage of state estimation is

that observations and modeling are combined, leading to a more accurate description of the oceanic system with insight into the underlying processes. ECCOv4r2 is used for all experiments completed in this study (Forget et al., 2016). ECCOv4r2 covers the time domain from 1992 to 2012. A lat-lon-cap grid is used with a varying resolution of around 100 km at the mid-latitudes and 40 km at the poles (Forget, Campin, et al., 2015).

In this configuration, the air-sea heat fluxes are prescribed using bulk formulas which act as a restoring term (Large & Yeager, 2004), acting to return the surface ocean back to an equilibrium with the overlying atmosphere. There are many options for prescribing surface boundary conditions in ocean models. One choice is to impose a surface heat flux, which would lead to no effective damping of surface temperature anomalies. Another choice is to impose the surface boundary conditions (for 2 m air temperature, 2 m specific humidity, wind speed, downward shortwave and longwave radiation) which are used to solve for the surface heat flux through the use of bulk formulas (Jones et al., 2019). This choice of boundary condition provides an effective damping of surface temperature anomalies (Williams, 1988). This choice is applied in the ECCOv4r2 set up that is used within these experiments. So, the effective atmospheric damping is not explicit, but is used to describe the effective response of the air-sea exchange and the model closure. The alternative choice is to employ a coupled model that solves for atmospheric and ocean temperature, humidity, and winds, and thus solves explicitly for surface fluxes. This choice also provides an effective damping of surface temperature, but it is not as strong as in the bulk formulas case (and avoids the atmosphere providing an unrealistically large heat source or sink to the ocean).

There is good agreement between the ECCOv4r2 state estimate and Argo observations over the upper ocean potential temperature and salinity in Southern Ocean mode water formation sites (Boland et al., 2021) and in the northern hemisphere (Jones et al., 2018). There is also good agreement between ECCO mode water properties in the South Indian and Southeast Pacific and satellite observations (Jones et al., 2019) and ECCOv4r2 closely agrees with the geography and magnitude of observed mixed layers (Forget, Ferreira, & Liang, 2015).

2.2. Adjoint Approach

In traditional model approaches, forward perturbation experiments define a quantity of interest and chosen area, which is then perturbed by a small amount. The effects of the perturbation on the model state are seen in differences from the unperturbed control state. For adjoint sensitivity experiments one quantity of interest, or objective function, is chosen and then the sensitivities of this function to every independent variable at each time, latitude, longitude, and depth are calculated backwards in time. The adjoint sensitivity experiments are used to span parameter space, and time, latitude, longitude, and depth space, in a computationally efficient manner. Hence, adjoint sensitivity experiments can be more computationally efficient than completing many instances of the more standard forward perturbation experiments. To obtain a comparable amount of information from the forward perturbation experiments when compared to adjoint sensitivity experiments would take an unfeasibly large number of experiments, much more computational time, and much more user time. The output of the adjoint sensitivity experiments are linear gradients, or sensitivities, of the objective functions to independent variables, which can be used to identify potential causal mechanisms. The sensitivities produced are linear approximations, however for large-scale oceanography problems, with objective functions averaged over years to months and large areas, the linear approximation is typically valid for several years to decades (Jones et al., 2018). Adjoint sensitivity experiments have been applied to understand ocean structures and variability in many areas of the global ocean (Fukumori et al., 2007; Smith & Heimbach, 2019).

2.2.1. Potential Vorticity Analysis

A limitation of the ECCO state estimate is that there are a limited number of objective functions available in the existing model code, and large-scale PV is not one of them. To overcome this limitation, we employ the equation of state to combine the objective functions of potential temperature, θ, and salinity, S, over two different depth regions to define the desired objective function of large-scale PV. Large-scale PV is defined as:

$$PV = \frac{-f}{\rho}\frac{\partial \sigma}{\partial z}, \tag{1}$$

where f is the Coriolis parameter, ρ is a reference ocean density, and σ is potential density. Large-scale PV is a measure of stratification: low PV refers to weak stratification, that is, a state wherein density does not change

rapidly with depth, whereas high PV refers to strong stratification. The overall objective function for the adjoint sensitivity experiments of large-scale PV is defined using four separate objective functions as follows;

$$J_1(\theta_u) = \frac{1}{V_u \Delta t} \int_{t_1}^{t_2} \int_{z_{u2}}^{z_{u1}} \int_A \int \theta \quad dx\, dy\, dz dt, \tag{2}$$

$$J_2(\theta_l) = \frac{1}{V_l \Delta t} \int_{t_1}^{t_2} \int_{z_{l2}}^{z_{l1}} \int_A \int \theta \quad dx\, dy\, dz dt, \tag{3}$$

$$J_3(S_u) = \frac{1}{V_u \Delta t} \int_{t_1}^{t_2} \int_{z_{u2}}^{z_{u1}} \int_A \int S \quad dx\, dy\, dz dt, \tag{4}$$

$$J_4(S_l) = \frac{1}{V_l \Delta t} \int_{t_1}^{t_2} \int_{z_{l2}}^{z_{l1}} \int_A \int S \quad dx\, dy\, dz dt, \tag{5}$$

where V_u and V_l represent the chosen upper and lower control volumes respectively, Δt is the time interval, t_1 and t_2 are the start and end dates of the time interval, z_{u1}, and z_{u2}, are the upper bounding depth levels, z_{l1}, and z_{l2}, are the lower bounding depth levels, and A represents the chosen latitude and longitude area. Each objective function represents the annual-average potential temperature, J_1, J_2, or salinity, J_3, J_4, over the chosen control volume and time interval. The final objective function, J, for average overall PV over a horizontal control volume, vertical control volume, and time interval is given by;

$$J = \frac{-f}{\Delta z}(-\alpha(J_1(\theta_u) - J_2(\theta_l)) + \beta(J_3(S_u) - J_4(S_l))), \tag{6}$$

where Δz represents the change over depth. The thermal expansion coefficient, α, measures the amount that a water parcel expands depending on temperature changes. The haline contraction coefficient, β, measures the amount that a water parcel contracts depending on salinity changes.

The adjoint sensitivities from the four separate adjoint sensitivity experiments are combined, via the equation of state, to obtain the sensitivity of large-scale PV to any of the independent variables,

$$\frac{\partial J}{\partial \mathbf{C}} = \frac{2f}{((z_{l2} + z_{l1}) - (z_{u2} + z_{u1}))}\left(-\alpha\left(\frac{\partial(J_1(\theta_u))}{\partial \mathbf{C}} - \frac{\partial(J_2(\theta_l))}{\partial \mathbf{C}}\right) + \beta\left(\frac{\partial(J_3(S_u))}{\partial \mathbf{C}} - \frac{\partial(J_4(S_l))}{\partial \mathbf{C}}\right)\right), \tag{7}$$

where $\mathbf{C}$ represents any independent variable, such as surface heat flux.

The chosen horizontal control volume is an area of deep winter mixed layers in the South Indian Ocean, which is a known area of mode water formation (Figure 2a). The first vertical control volume extends from the surface to 189 m and the second vertical control volume extends from 444 to 657 m (Figure 2b). These depths are chosen to examine the physical controls of the large-scale stratification in mode water formation regions. We are interested in the overall subduction and ventilation pathways and processes, which involve mode water formation as an important component, but extend over broader depth and density ranges (Figure 1). This construction allows us to investigate the controls on large-scale stratification over the top several hundred meters of the ocean. Several depth ranges were investigated, and the result was not found to be sensitive to exact choices, provided there was enough separation between the upper and lower depths and that they captured the essential qualities. The time interval chosen is the annual-average of large-scale PV over year 2011 (Figure 1), as we are interested in seeing how the mode water is forced over several annual cycles. This construction is similar to other adjoint studies investigating the forcing of mode water properties (Jones et al., 2020). Several time intervals were tried in years with differing atmospheric forcing and the results were found to be insensitive to changes in this choice.

Another way that adjoint sensitivities can be analyzed is by convolving the sensitivities with forcing fields to produce impacts (Jones et al., 2018). Impacts take into account variability in the forcing. Impacts of large-scale PV to surface heat flux, I_H, are calculated such that:

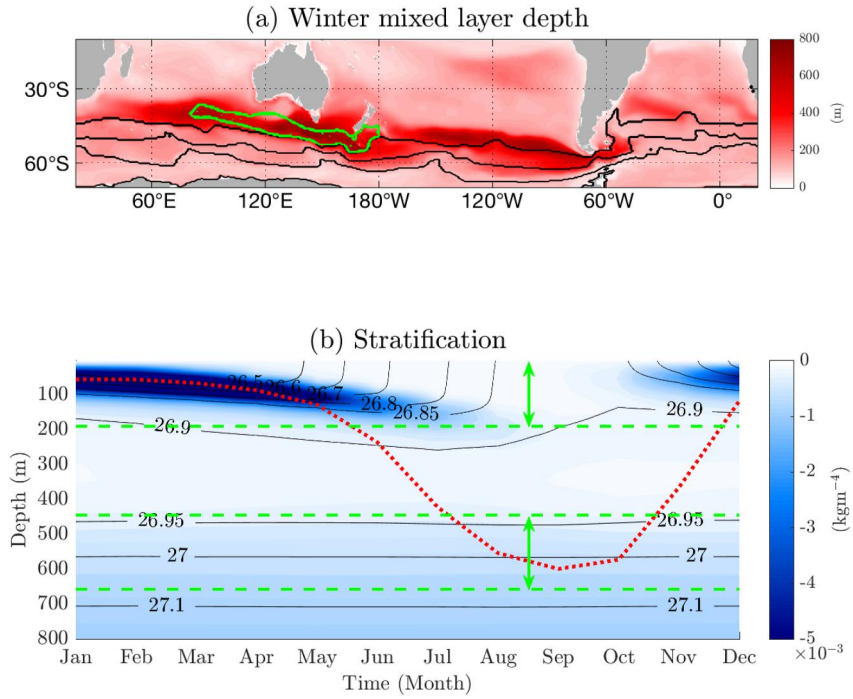

Figure 2. (a) Winter (July, August, September) mean mixed layer depth (m, red) together with sea surface height contours (−0.25, −1, −1.5 m, black) defining the fronts of the Antarctic Circumpolar Current and the South Indian adjoint sensitivity target region (green contour). (b) Stratification (kgm^{-4}, blue) in the South Indian Ocean target region together with neutral density surfaces (kgm^{-3}, black) and mixed layer depth (m, red). Green dashed lines show upper and lower vertical regions used in the adjoint sensitivity experiments for density.

$$I_{\mathcal{H}} = \frac{\partial PV}{\partial \mathcal{H}} \times \left(\mathcal{H} - \overline{\mathcal{H}}\right), \tag{8}$$

where $\mathcal{H}$ represents surface heat flux and $\left(\mathcal{H} - \overline{\mathcal{H}}\right)$ represents surface heat flux anomaly relative to the annual average which is calculated including the seasonal cycle.

Both sensitivities and impacts are useful, providing complementary information as to the intrinsic sensitivity of a property and the extent that surface forcing affects that property.

2.3. Forward Perturbation Approach

Forward perturbation experiments are used to illustrate the local and non-local effects of surface fluxes on large-scale PV in mode water formation regions. Forward perturbation experiments are completed by perturbing one variable of interest and then integrating the model to see how this changes other variables. First, a control experiment is integrated with no model changes. Then, the perturbation experiment is integrated which contains the information for the imposed perturbation. To see how the perturbation affects the model state the difference between the perturbation and control experiment is taken.

The perturbation experiments discussed in Section 4.2 consider a surface heat flux perturbation in a region upstream of the adjoint sensitivity experiment control volume within the ACC. This heat flux perturbation experiment is conducted to confirm which physical mechanisms dominate the response and assess whether the response is in accord with the adjoint sensitivities of the large-scale PV to surface heat flux. The horizontal area that the surface heat flux perturbation is imposed over is chosen to be an area where strong positive sensitivities of stratification to surface heat flux are seen in Figure S1 in Supporting Information S1. The surface heat flux perturbation of magnitude 100 W m^{-2} is imposed for 3 months. The results of this perturbation experiment are discussed using the perturbation minus the control, to understand how the perturbation alters the model state.

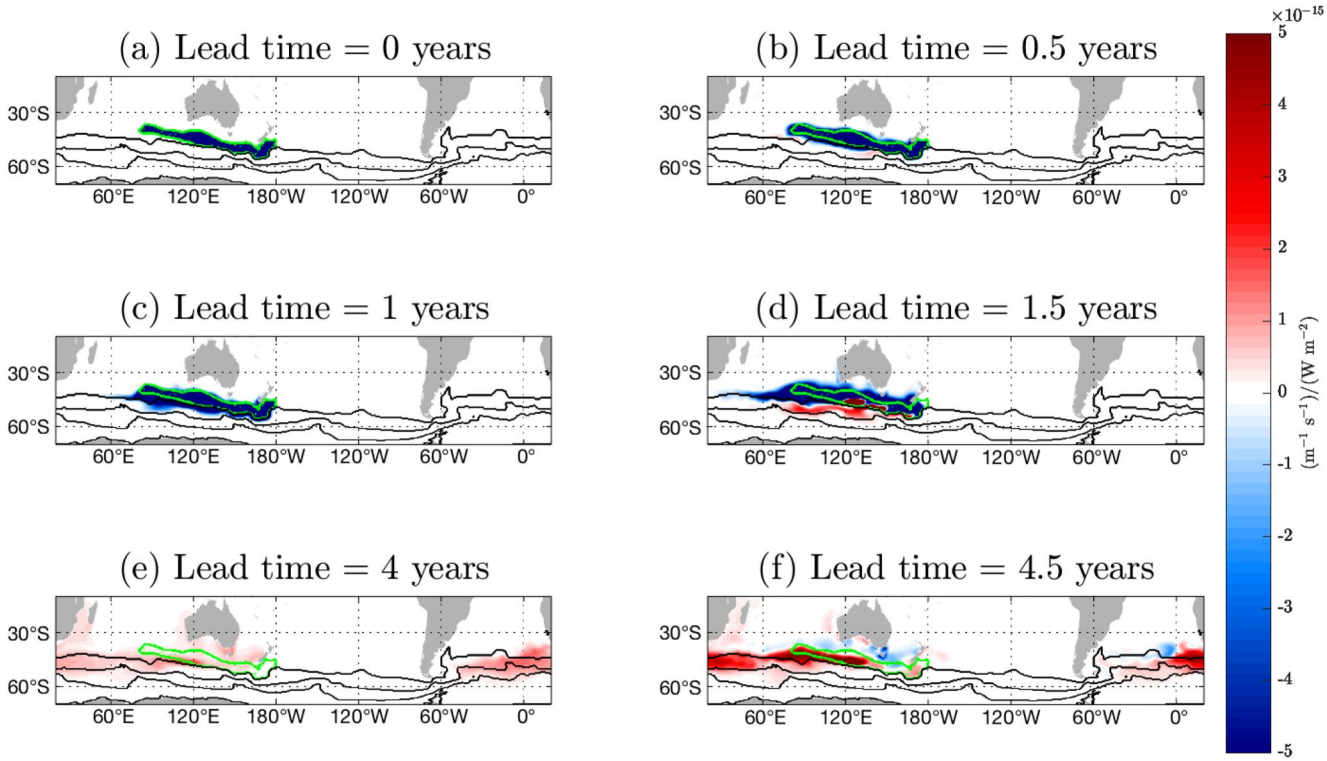

Figure 3. Sensitivity of large-scale potential vorticity to surface heat flux ($(m^{-1}\ s^{-1})/(W\ m^{-2})$) at the surface for lead times for 0, 0.5, 1, 1.5, 4, and 4.5 years. The South Indian adjoint sensitivity target region (green) and frontal zones of the Antarctic Circumpolar Current (black contours, as in Figure 2) are included.

3. Adjoint Sensitivity Experiment Analysis

3.1. Sensitivity of Large-Scale PV to Surface Heat Flux

First, the adjoint sensitivities of large-scale PV to surface heat flux, $\frac{\partial J}{\partial \mathcal{H}}$ (Equation 7), are considered. The sign of the adjoint sensitivities is important for the interpretation of the physical mechanisms that drive changes in large-scale PV or equivalently stratification.

Positive surface heat flux is defined to be out of the ocean leading to a surface cooling, while negative surface heat flux is defined to be into the ocean leading to surface warming. This definition is consistent with the ECCOv4r2 state estimate. Positive sensitivities indicate that if there is a positive surface heat flux anomaly, leading to surface cooling, then there will be an increase in PV, as well as indicating if there is a negative surface heat flux anomaly, leading to surface warming, then there is a decrease in PV. Conversely, negative sensitivities indicate that if surface cooling occurs then there will be a decrease in PV in the target region. Also, negative sensitivities indicate that if there is a surface warming then there will be an increase in PV in the target region. The averaged integral function, or objective function J, is imposed over 0–1 year lead time, where 0 is defined as the start of the target year in the objective function, December 2011, and 1 as the end of the target year, January 2011. So, lead time is defined as the measure of time since the objective function has been imposed. The initial adjoint sensitivities are calculated over the year where the objective function is present.

Initially, the sensitivities of large-scale PV to surface heat flux are mostly negative and local to the control volume region (Figures 3a and 3b). This response to local cooling is intuitively expected to lead to a weakening of stratification through local convection. After 1.5 years lead time, positive sensitivities are found locally and to the south of the control volume (Figure 3d). Over lead times from 2 to 4.5 years, the positive sensitivities dominate and spread westerly due to ACC advection (Figures 3d and 3e); the full temporal evolution is shown in Figure S1 in Supporting Information S1. This positive sensitivity is unexpected and reveals that on longer lead times surface cooling leads to a strengthening of stratification; this counter-intuitive response is explored in more detail in Section 4.

3.2. Sensitivity of Large-Scale PV to Other Independent Variables

The sensitivities of large-scale PV to surface freshwater flux, $\frac{\partial J}{\partial FW}$ (Equation 7), are mainly negative on all linear lead times (Figure S2 in Supporting Information S1), implying that if there is a positive anomaly from a freshwater flux out of the ocean to the atmosphere, leading to an increase in surface salinity, then there will be a consequent decrease in stratification. This response is consistent with increased evaporation leading to an increase in surface density and convection, and therefore ultimately a decrease in stratification.

The sensitivity of large-scale PV to eastward wind stress, $\frac{\partial J}{\partial \tau_E}$ (Equation 7), initially shows a complicated local pattern of both positive and negative sensitivities. On longer lead times, a dipole forms with positive sensitivity to the north of the control volume and negative sensitivity locally and to the south of the control volume (Figure S3 in Supporting Information S1). Dipoles in wind-stress sensitivities indicate regions where the wind stress alters horizontal Ekman transport and its curl, which drives vertical movement of the density surfaces. In the Southern Ocean, an eastward wind stress leads to northward Ekman transport, which tilts the density surfaces upwards to the south and downwards to the north. Jones et al. (2018, Figure 7) shows how the opposing sign of eastward and westward wind stress perturbations lead to changes in density surface position. A positive eastward wind stress in the Southern Ocean leads to a northward Ekman transport, which over a northern region of positive sensitivities (Figures S3e–S3l in Supporting Information S1) indicate a strengthening in stratification within the control volume. At the same time, a stronger eastward stress over the southern region of negative sensitivity indicate a weakening of stratification within the control volume. So, these two contributions are acting in competition to each other over the control volume, when the same sign forcing is applied to both. There is also a signal of westward propagation of Rossby waves over the Pacific basin.

The sensitivities of large-scale PV to northward wind stress, $\frac{\partial J}{\partial \tau_N}$ (Equation 7), are stripes of both positive and negative sign, but without a coherent pattern emerging as for zonal wind stress (Figure S4 in Supporting Information S1). On lead times longer than 1 year there is a signal of communication via an eastern boundary wave along the coast of South America, as has been noted in other studies of the Southern Ocean mixed layer (Boland et al., 2021).

3.3. Heat Flux Is the Dominant Impact on Stratification

The sensitivities of large-scale PV to surface heat flux, surface freshwater flux, eastward wind stress, and northward wind stress are now compared to each other using impacts, the product of the sensitivity and forcing variability. As impacts take into account forcing variability, separate impacts can be directly compared and contrasted. Impacts, $I_{\mathbf{C}}$, are found through multiplying adjoint sensitivities by the variability of the respective forcing fields (e.g., Equation 8 for surface heat flux). The impacts of surface heat flux are at least one order of magnitude larger than the impacts of the other independent variables (Figure 4a). On lead times of 2–3 years, eastward wind stress is also important in affecting the formation of mode water in the South Indian region (Figure 4b), becoming of comparable importance to upstream heat fluxes in contributing to the stratification changes. Eastward wind stress leads to northward Ekman transport in the Southern Ocean, which then tilts the density surfaces and alters the stratification of the control volume. The effects of surface freshwater flux and northward wind stress on stratification are relatively small, $\mathcal{O}\left(10^{-15}\right)$, for all lead times.

This impact analysis reveals that the impacts of surface heat flux are dominant over other forcing factors in setting stratification of mode water formation regions. Therefore, the next section focuses on understanding the temporal evolution of the sensitivities of large-scale PV to surface heat flux.

4. The Time Dependence of the Sensitivity of Large-Scale PV to Surface Heat Flux

The physical processes leading to the sign change in the sensitivity of large-scale PV to surface heat flux, illustrated in Figure 3, are now explored in more detail.

4.1. Adjoint Sensitivities Split Into Density Component Contributions

To consider the time dependence of the adjoint sensitivities, Hovmöller diagrams are used to isolate the mechanisms that lead to the change of sign over lead time for the sensitivity of large-scale PV to surface heat flux. Due

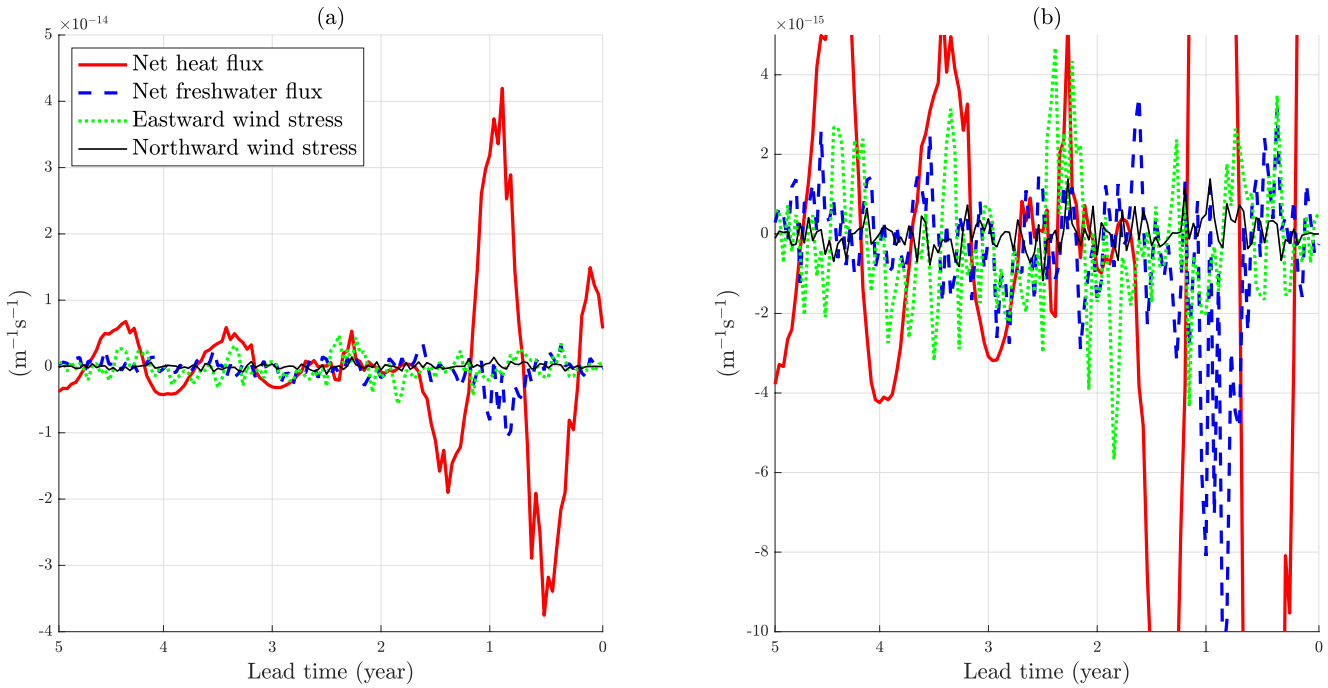

Figure 4. (a) Time series over lead time (year) of the total impacts of surface heat flux (sensitivity of large-scale potential vorticity to surface heat flux in the South Indian Ocean multiplied by heat flux anomalies) (red thick, $m^{-1}\ s^{-1}$), surface freshwater flux (blue dashed, $m^{-1}\ s^{-1}$), eastward wind stress (green dotted, $m^{-1}\ s^{-1}$), and northward wind stress (black thin, $m^{-1}\ s^{-1}$). (b) Same as (a) with small y axis to show more detail.

to the construction of the sensitivities of large-scale PV (Equation 7), the sensitivities can be separated into the component parts of upper thermal ($J_1(\theta_u)$), upper haline ($J_2(S_u)$), lower thermal ($J_3(\theta_l)$), and lower haline ($J_4(S_l)$).

On lead times longer than 2 years, there is a change of sign from mainly negative sensitivities to mainly positive sensitivities for large-scale PV to surface heat flux (Figure 5a). The upper depth density contribution to this signal is very strong on lead times less than 2 years, exhibiting both negative and positive sensitivities, but their sensitivities decreases in strength on longer lead times (Figure 5b). The lower density contribution remains consistently at a similar order of magnitude after 1 year lead time (Figure 5c). So, the change of sign in adjoint sensitivity over time is due to the upper density contribution.

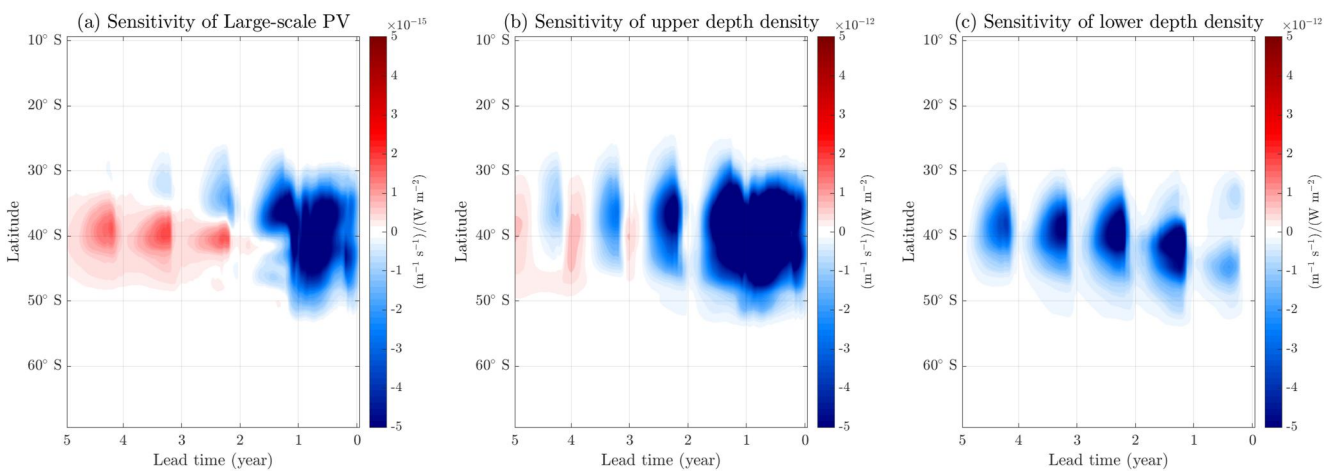

Figure 5. (a) Hovmöller figure for the sensitivity of large-scale potential vorticity to surface heat flux in the South Indian Ocean ($(m^{-1}\ s^{-1})/(W\ m^{-2})$) for lead time (year) versus latitude (°). Likewise sensitivities of density to surface heat flux for the (b) upper (0–189 m) and (c) lower (444–657 m) depths ($(m^{-1}\ s^{-1})/(W\ m^{-2})$).

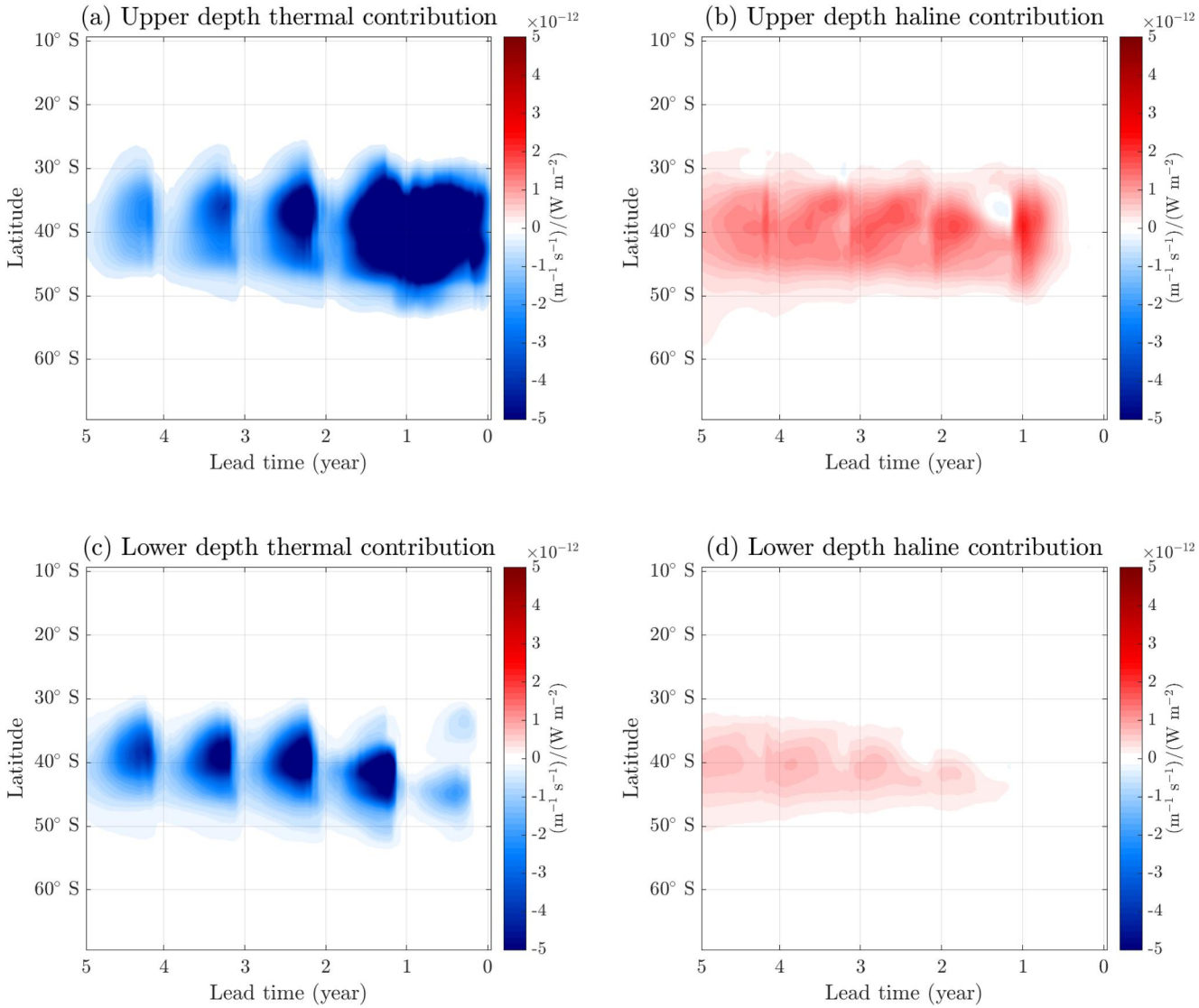

Figure 6. (a) Hovmöller figure for the zonally-integrated upper depth thermal contribution to density in the South Indian Ocean ($(m^{-1}\ s^{-1})/(W\ m^{-2})$) for lead time (year) verses latitude (°). Likewise for the (b) zonally-integrated upper depth haline contribution to density, (c) zonally-integrated lower depth thermal contribution to density, and (d) zonally-integrated lower depth haline contribution to density ($(m^{-1}\ s^{-1})/(W\ m^{-2})$).

To further isolate how the change of sign occurs, the sensitivities are separated into upper and lower thermal, $J_1(\theta_u)$, $J_3(\theta_l)$, and haline, $J_2(S_u)$, $J_4(S_l)$, contributions. The upper depth thermal contribution is negative on all lead times, but decreases in time (Figure 6a). The upper depth haline contribution is positive on all lead times (Figure 6b). When both the upper depth thermal and haline contributions are combined in terms of their effect on density, the change of sign in the sensitivity of the upper density comes from the thermal contribution weakening over time, whilst the haline contribution remains at the same magnitude in time. The lower depth thermal contribution (Figure 6c) is dominant over the lower depth haline contribution (Figure 6d) on all lead times.

The upper ocean thermal contribution weakens in time (Figure 6a), in contrast to the more constant order of magnitude of the upper ocean haline contribution (Figure 6b) and both the deeper thermal and haline contributions (Figures 6c and 6d). The weakening of the upper thermal contribution in time is in accord with the effect of air-sea flux exchanges with the atmosphere, acting to effectively dampen the upper ocean temperature contribution (Williams, 1988). Hence, the change of sign in the sensitivity of stratification to surface heat flux is viewed as due to an initial surface thermal response being replaced by a delayed haline response.

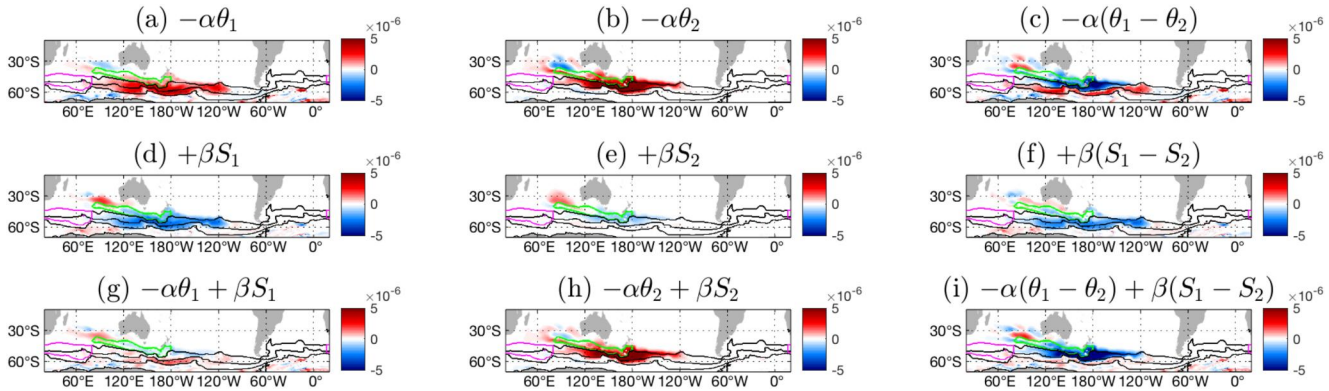

Figure 7. Upstream surface cooling perturbation experiment minus control experiment after 3 years for each of the component parts of large-scale potential vorticity. (a) Upper depth thermal component, (b) lower depth thermal component, (c) upper depth minus lower depth thermal component, (d) upper depth haline component, (e) lower depth haline component, (f) upper depth minus lower depth haline component, (g) upper depth density component, (h) lower depth density component, and (i) upper depth minus lower depth density component. Together with sea surface height contours (−0.25, −1, −1.5 m, black) defining the fronts of the Antarctic Circumpolar Current, the South Indian adjoint sensitivity target region (green contour), and the region where the 100 Wm^{-2} heat flux perturbation is imposed for 3 months in winter (magenta contour).

4.2. Forward Perturbation Experiment Analysis

Forward perturbation experiments are used to help understand the physical mechanisms that lead to the change of sign in the adjoint sensitivity of large-scale PV to surface heat flux discussed in the previous section; this combination of adjoint analyses and forward experiments is increasingly common (Boland et al., 2021; Jones et al., 2019; Marotzke et al., 1999; Verdy et al., 2014), although there are other methods, such as using adjoint sensitivities, to reconstruct the forward model state (Smith & Heimbach, 2019). These forward perturbation experiments are used as the adjoint sensitivities only identify causal mechanisms within the model, but do not explain physically how the variables are related. The same forward perturbation experiments are used to identify how long the linear approximation of the adjoint sensitivities is valid for (Appendix A).

4.2.1. Potential Temperature, Salinity and Stratification Responses to Surface Heat Flux

To gain insight into the evolution of the upper ocean temperature and salinity contributions to density and stratification, a forward perturbation experiment is conducted. A surface heat flux perturbation of 100 W m^{-2} is imposed for 3 months in an upstream region located within the ACC (Figure 7, magenta contour). This location is consistent with the central location of positive sensitivities at approximately 2–3 years lead time (Figures S1e and S1g in Supporting Information S1). The response of the density anomaly produced by the perturbation is considered in terms of thermal and haline contributions.

The initial response to the upstream cooling perturbation is dominated by the upper thermal contribution (Figure S5i in Supporting Information S1). Cooling leads to an increase in the thermal contribution and a smaller increase in the haline contribution (Figure S5 in Supporting Information S1).

However, following this initial response, there is a delayed response (Figure 7), with the surface cooling perturbation leading to a weakening of stratification within the South Indian control volume (Figure 7i). This stratification response mainly comes from the upper haline contribution (Figure 7d). Also, in some small local areas, and just outside of the control volume, a weakening of stratification (Figure 7i) comes from the lower depth thermal contribution (Figure 7b) becoming larger than the upper depth thermal contribution at timescales beyond 2–3 years (Figure 7a), highlighting the importance of advection and subduction into the mode water region and the role that effective atmospheric damping plays on surface heat perturbations.

Previously, only the surface response to heat flux was considered for each of the four objective functions ($J_1(\theta_u)$, $J_2(\theta_l)$, $J_3(S_u)$, $J_4(S_l)$) which are defined at different depths. Now, the response over depth is also considered in order to understand the dynamics at play within the mode water formation site response to surface heat fluxes (Figure 8). To examine the evolution of the anomaly as it progresses downstream from the site of the initial perturbation to the mode water formation site, we consider the depth structure of the anomaly in a quasi-Lagrangian fashion (Figure 8).

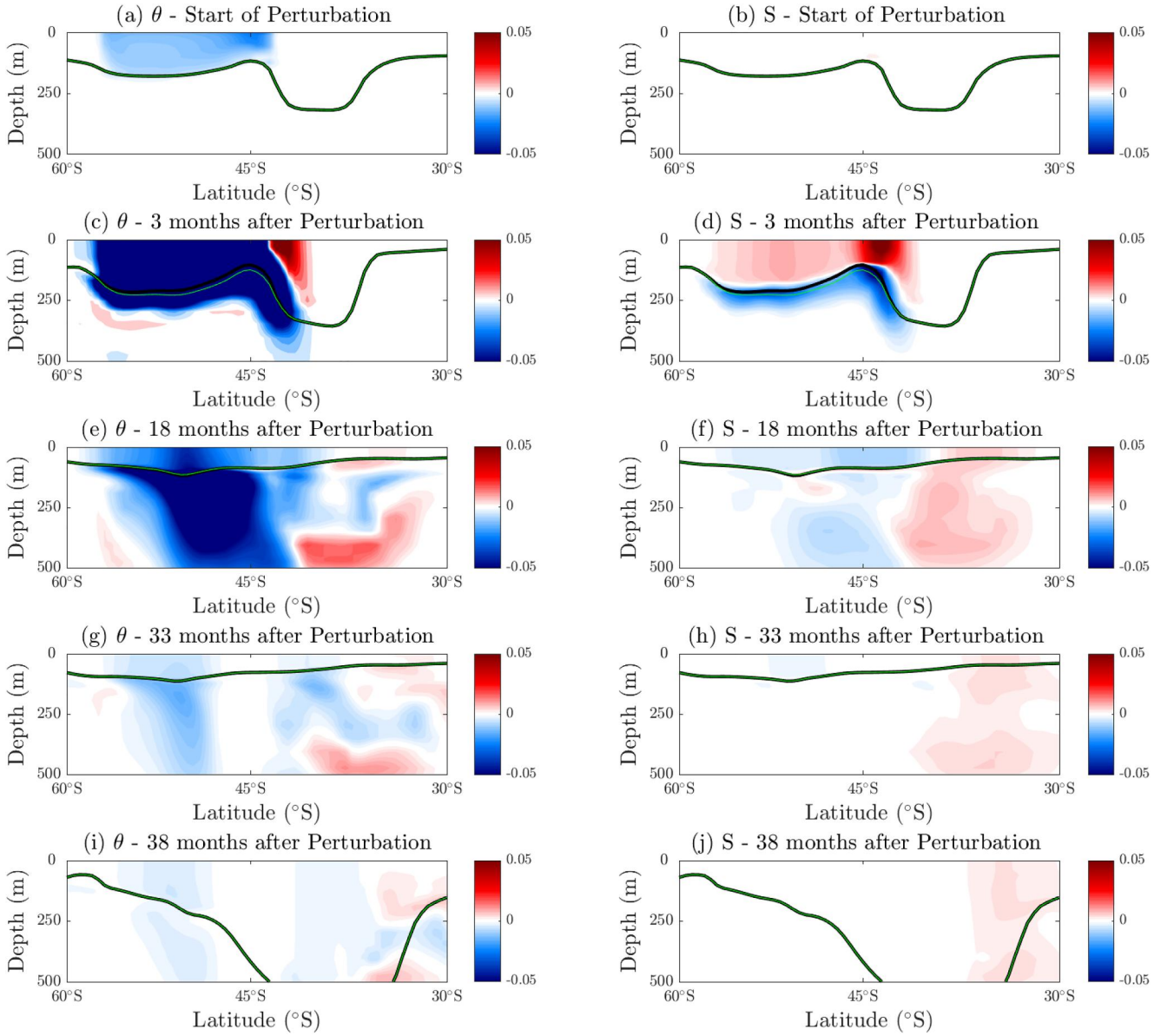

Figure 8. Upstream cooling perturbation experiment minus control experiment for potential temperature (°C) averaged over longitudes 55–65°E (a, c), and averaged over longitudes 90–110°E (e, g, i). Likewise for salinity (PSU) (b, d, f, h, j). Together with perturbation mixed layer depth (green) and control mixed layer depth (black). The top two rows show potential temperature and salinity responses (Figures 8a–8d) at the longitudes of the perturbation site, approximately 60°E. The bottom three rows (Figures 8e–8j) follow the responses to the heat flux perturbation as they are advected eastward via the Antarctic Circumpolar Current, approximately 90°E–110°E.

The initial surface heat loss leads from the upstream cooling anomaly is within the mixed layer, which is local to where the perturbation is applied (Figure 8a). This decrease in potential temperature is local to the region of the perturbation, and does not yet affect the South Indian mode water region. Subsequently, over the next 3 months, the cool surface anomaly spreads into the thermocline via subduction through the base of the mixed layer (Figure 8c). Over the following year, the temperature anomaly weakens within the mixed layer (Figure 8e). After approximately 3 years, the response becomes even weaker, with most of the cooling anomaly being further south than the control volume (Figures 8g and 8i).

The surface heat flux loss does not initially lead to a salinity anomaly (Figure 8b). However, after 3 months, there is a thickening of the mixed layer and an enhanced entrainment from the thermocline to the mixed layer (Figure 8d, with the green line deeper than black line). This enhanced entrainment then transfers salt from within the thermocline to the mixed layer, leading an increase in salinity within the mixed layer and a decrease in the

thermocline (Figure 8d). These salinity anomalies are then advected and mixed downstream into the control volume. After approximately 3 years, the salinity response to an upstream cooling perturbation within the latitudes of the South Indian control volume is an increase, which is the dominant response on these timescales (Figures 8h and 8j). This dominant salinity response is of the opposite sign to the initial dominant thermal response. These contrasting responses highlight the importance of advection in the salinity response and effective atmospheric damping in the surface potential temperature response.

The mixed layer response to perturbations is important in modulating the local variations in the thermal and haline responses (Figures 8c and 8d). However, the mixed layer depth changes are not transferred along the streamlines of the ACC, but remain within the region that the surface heat flux perturbation is imposed (Figures 8e–8h). This mixed layer response may be exacerbated due to the fact that the surface heat flux perturbation is imposed in austral winter, when there are deep mixed layers in mode water formation regions. As the resulting upper ocean anomalies are advected downstream via the ACC over time the mixed layer depth in the mode water formation region has seasonally shoaled. So, the mixed layer depth over the mode water formation region is not strongly influenced by an upstream perturbation.

Hence, the forward perturbation experiments reveal that a surface heat flux anomaly creates a temperature anomaly within the mixed layer and thermocline, that is preferentially damped in the mixed layer. The surface heat flux anomaly indirectly creates a salinity anomaly by initially altering the vertical transfer of salinity through entrainment. Then these salinity anomalies are advected downstream. This contribution leads to the two-phase response of mode water region stratification to surface heat fluxes. Strong flows are a requirement for this two-phase response to a heat flux perturbation as the response is rapidly advected downstream, away from the local perturbation region.

5. Discussion

The effect of surface heat fluxes is examined on the mode waters in the South Indian sector of the Southern Ocean. Initially, surface cooling leads to a local decrease in temperature and an associated weakening of the stratification. Then, on lead times of over 2 years, there is a lagged response involving an effective atmospheric damping of the surface thermal response and an opposing advective haline response, leading to stronger stratification. The effective atmospheric damping effect is particularly strong due to the bulk formulas used in this model setup, which uses prescribed atmospheric temperatures. The air-sea heat fluxes act as a restoring term, returning the surface ocean back toward equilibrium with the overlying atmosphere. Due to this restoring, our results should be interpreted in the context of the limiting case of strong effective damping. By contrast, in a setup with prescribed air-sea heat fluxes, there would be very weak to no local effective atmospheric damping. A coupled model is expected to behave between these two extreme cases; the effects of different model setups on adjoint sensitivities is explored further by Czeschel et al. (2010), Kostov et al. (2019), and Pillar et al. (2016).

The extent that the two-phase response in the sensitivity of stratification to surface heat flux carries over to other Southern Ocean mode water formation regions is next assessed by considering the Southeast Pacific. The sensitivities of large-scale PV in the Southeast Pacific mode water formation to surface heat flux is found to be very similar to the South Indian mode water formation region results (Figure 9 and Figure S6 in Supporting Information S1). In both the South Indian and Southeast Pacific mode water formation regions, winter surface cooling is a key process in mode water formation (Tamsitt et al., 2020). Turbulent heat loss events are particularly important, as shown from surface flux mooring data (Ogle et al., 2018). The mode water responses differ interannually in both regions, with central and eastern pool thicknesses varying out of phase with each other (Cerovečki & Meijers, 2021).

While there are differences in the exact magnitude and timing of the response to surface forcing between the two regions, the same fundamental two-phase response of stratification occurs in both the Southeast Pacific and the South Indian regions. The surface heat flux again leads to contrasting temporal anomalies in stratification in the Southeast Pacific section (Figure 10b); there is a change in sign in the sensitivity of large-scale PV to surface heat flux from negative to positive on lead times of 2 years. The timescales of the change of sign in sensitivity are slightly different in both regions, with the positive sensitivities in the Southeast Pacific remaining to the south of the control volume for longer than in the South Indian. This difference could be due to the different latitudinal structure of each region, with the fronts of the ACC intersecting more closely with the Southeast Pacific region

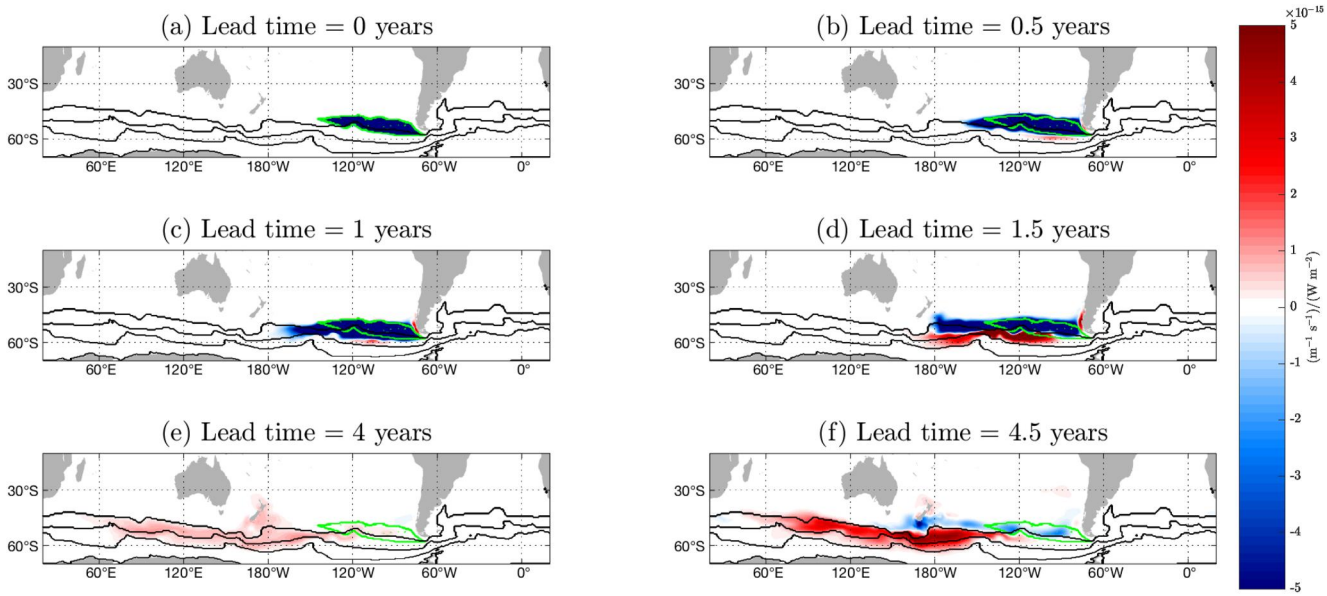

Figure 9. Sensitivity of large-scale potential vorticity to surface heat flux ($(m^{-1}\ s^{-1})/(W\ m^{-2})$) at the surface for lead times for 0, 0.5, 1, 1.5, 4, and 4.5 years. The Southeast Pacific adjoint sensitivity target region (green) and frontal zones of the Antarctic Circumpolar Current (black contours, as in Figure 2) are included.

and leading to more rapid advection. Local negative sensitivities in winter remain on longer lead times, up to 5 years, in the Southeast Pacific than in the South Indian. This similar response is probably due to both regions being similarly connected to along stream flows in the ACC, and hence advection into both of these areas is important, as is meridional transport of surface signals via Ekman fluxes.

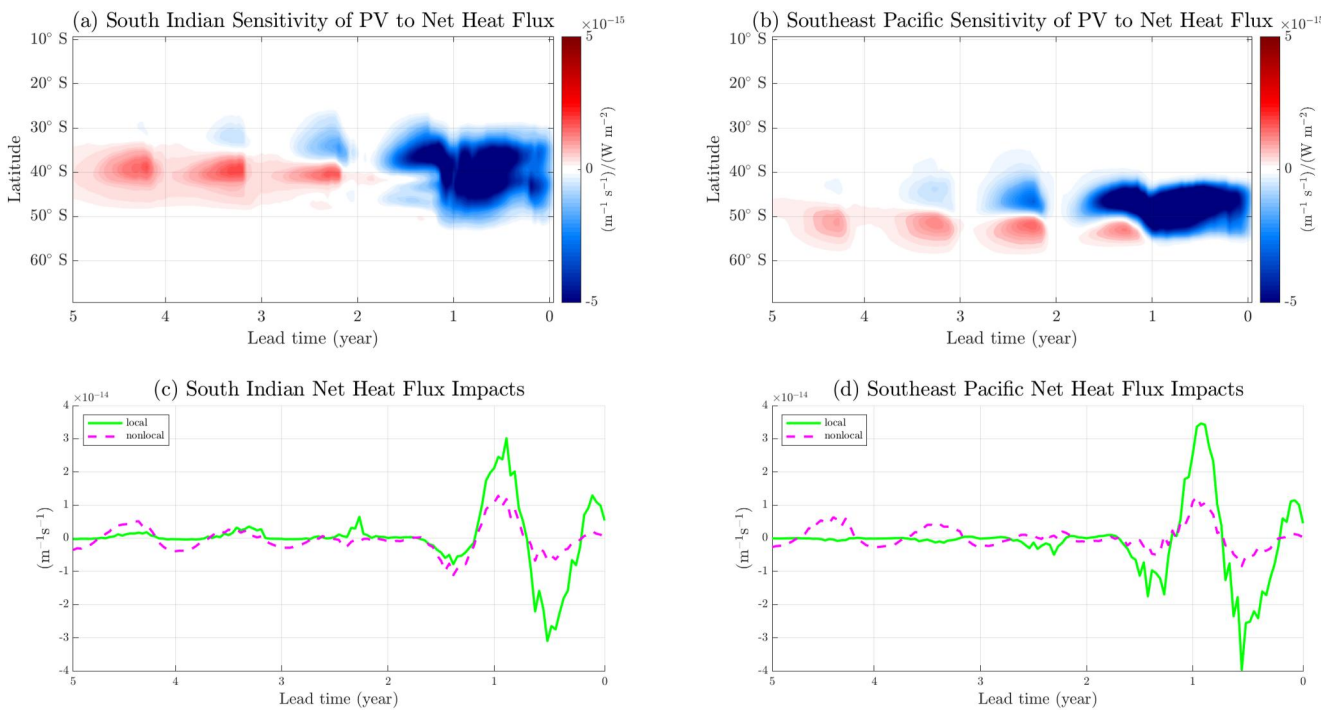

Figure 10. (a, b) Hovmöller figure for the sensitivity of large-scale potential vorticity (PV) to surface heat flux in the South Indian Ocean (left panel) and the Southeast Pacific Ocean (right panel) ($(m^{-1}\ s^{-1})/(W\ m^{-2})$) for lead time (year) versus latitude (°). (c, d) Time series over lead time (year) of the impacts of surface heat flux (sensitivity of large-scale PV to surface heat flux multiplied by heat flux anomalies) in the South Indian Ocean (left panel) and the Southeast Pacific Ocean (right panel) separated into local (green) and non-local (magenta) components ($m^{-1}\ s^{-1}$).

The mode water response to surface forcing involves both local and remote contributions (Figures 10c and 10d). Surface heat flux is the most important impact on stratification in mode water formation regions (Figure 4a). On lead times of up to 1 year, the objective function year, in both mode water formation regions the local impacts dominate with small remote impacts. The seasonality of the surface heat flux anomalies leads to the negative sensitivities becoming positive and negative impacts depending on season. On longer lead times, the advective contribution becomes increasingly important and the local contribution is much weaker. Consistent with our diagnostics, Cerovečki and Meijers (2021) and Gao et al. (2018) also find that upstream forcing is important in the formation of mode waters.

These signals are due to the adjoint sensitivities measuring both the local and far-field contributions to a particular year, in this case evaluated for 2011. Thus, the mode water stratification in 2011 only has a limited memory for local forcing and upstream forcing is also important in determining its formation. In particular, the non-local impacts of heat flux become more important than the local impacts after 1.5 years lead time in the South Indian sector and after 2.5 years lead time in the Southeast Pacific sector. Hence, the memory of the ocean on local scales is much shorter than the upstream, advective contribution and upstream effects are communicated over time through advection by the currents of the ACC. In an analogous manner, the subpolar North Atlantic ocean heat content is viewed as being controlled by a combination of short-term, local and long-term, advective responses to atmospheric forcing, which have opposing signs (Khatri et al., 2022).

The sensitivity of the Atlantic Meridional Overturning Circulation (AMOC) to surface heat flux also exhibits a change of sign in adjoint sensitivity over time (Kostov et al., 2019). The sensitivities of the AMOC to surface heat flux have a seasonal change of sign (Kostov et al., 2019). The seasonal change of sign in sensitivity of AMOC to surface heat flux is not attributed to ocean circulation, but is instead attributed to mixed layer thermodynamics and local air-sea feedbacks (Kostov et al., 2019). In our analysis, the role of mixed layer processes are important in affecting temperature and salinity, but they do not provide the dominant control for stratification after 2 years lead time. In accord with our analysis, Kostov et al. (2019) finds an important role for effective atmospheric damping of potential temperature at the surface, as well as finding that changes in surface heat flux can impact salinity, as well as potential temperature, due to a strong coupling between the parameterized surface fluxes.

In the present study, the presence of thickness dipoles within both mode water formations regions (Cerovečki & Meijers, 2021; Meijers et al., 2019), has been largely ignored, and these dipoles may impact the sensitivities and forcing responses. However, Boland et al. (2021) considered multiple different years to assess the sensitivities of mixed layer depth and found there was not a wide spread in response.

The effect of surface heat flux in providing a two-phase response in stratification may be relevant in mode waters in other regions of the global ocean, such as the North Atlantic 18° water (Talley & Raymer, 1982). The longer timescale response involving advection depends on the strength and coherence of the flow, which is particularly strong in the ACC, and may carry over to boundary currents in ocean gyres.

6. Conclusion

Southern Ocean mode water formation regions are important in sequestering large amounts of heat and carbon, and their future responses to changes in winds and buoyancy fluxes are uncertain. Adjoint and forward model experiments are conducted to understand the processes that control the properties and stratification of these mode waters and how their response may change over time.

Adjoint sensitivity experiments are conducted in the two main mode water formation regions in the Southern Ocean, the South Indian and the Southeast Pacific, using an objective function defined by the annual-average stratification over the chosen control volume. The sensitivities of stratification to surface heat flux exhibit a change of sign from mostly negative to mostly positive on lead times over approximately 2 years. Furthermore, impacts, defined by sensitivities multiplied by respective forcing anomalies, reveal that surface heat flux is the most important impact on the stratification in mode water formation regions.

Due to the construction of the objective function, the sensitivities can be separated into upper and lower density contributions. This separation of the sensitivities reveals that the change of sign in the sensitivity of stratification comes from the upper density component, which can respond more rapidly to surface changes. Separating the upper density component further into thermal and haline components shows that over time the thermal contribution is effectively atmospherically damped and the haline contribution, initially driven by entertainment

triggered by the initial heat response and later involving an advective response, becomes dominant leading to an opposing-sign response. The strength of the effective atmospheric damping seen here affected by our choice of model setup. In a coupled model, the damping might be weaker leading to the advection of the haline contribution becoming dominant on longer lead times.

The two-phase response, which changes sign over time, holds under a traditional forward perturbation experimental framework. Initially, an imposed surface heat flux anomaly drives a surface temperature anomaly, which is subducted beneath the mixed layer. The potential temperature anomaly is subsequently effectively damped at the surface, becoming weak after approximately 2–3 years. The response of salinity is initially dominated by changes in entrainment from mixed layer deepening, but on longer timescales, the response of salinity to cooling is dominated by advection. The forward perturbation experiments also found that the linear approximation of the sensitivities is valid for approximately 5 years.

The effect of surface fluxes in driving a two-phase stratification response requires an advective regime where large potential temperature and salinity contrasts occur with depth. Therefore, further analysis of the two-phase response in different oceanic regimes is needed to reveal how representative this response is of the wider ocean. In addition, mode waters may be sensitive to different atmospheric regimes, as revealed in the different phases of the dipole in mixed layer depths in the Southeast Pacific and South Indian (Cerovečki & Meijers, 2021; Meijers et al., 2019).

Appendix A: Linearity of Sensitivities

Forward perturbation experiments, from Section 4.2, are used to identify how long the linear approximation of the adjoint sensitivities is valid for. This appendix follows the method of Verdy et al. (2014) to identify how long the adjoint sensitivities can be used to identify linear causal mechanisms.

Similar responses are found in potential temperature, salinity, and stratification after an identical upstream cooling or warming perturbation is imposed. However, the mixed layer response is of the opposite sign. The similar responses to identical positive and negative heat flux perturbations suggest there is a linear response on short timescales. The linear and non-linear components are assessed using the framework set out in Verdy et al. (2014), where an approximation of the linear terms is found from the difference between the positive and negative signed perturbations:

$$Linear = \frac{(\theta_+ - \theta_0) - (\theta_- - \theta_0)}{2}, \tag{A1}$$

here θ_+ represents the potential temperature with a positive heat flux perturbation, θ_0 represents the potential temperature in a control experiment, and θ_- represents the potential temperature with a negative heat flux perturbation. To find an approximation of the non-linear terms the sum of the positive and negative perturbations are taken, such that:

$$Non - linear = \frac{(\theta_+ - \theta_0) + (\theta_- - \theta_0)}{2}. \tag{A2}$$

Then the root mean square difference around the perturbation region is taken of the linear and non-linear responses. The forward perturbation experiments suggest that the linear approximation of the adjoint sensitivities is valid for approximately 5 years, before the nonlinear components become too large (Verdy et al., 2014).

Data Availability Statement

The data supporting our conclusions can be found at: https://zenodo.org/record/8160115 (Pimm, 2023). This site includes data from the four separate adjoint experiments that are combined to create the overall sensitivities of large-scale PV to surface heat flux (Equation 7) for 6 years lead time which supports our main conclusions. The ECCOv4r2 model setup used in this work is available for download on Github (https://github.com/gaelforget/ECCO_v4_r2) as an instance of the MIT general circulation model (MITgcm, http://mitgcm.org/). Adjoint code was generated using the TAF software tool, created and maintained by FastOpt GmbH (http://www.fastopt.com/).

Numerical model runs were carried out on ARCHER and ARCHER2, the UK national HPC facility (https://www.archer2.ac.uk/).

Acknowledgments

CP was supported by a UK NERC studentship and by a NERC Grant NE/T010657/1 (SARDINE). RW was supported by NERC Grants NE/T007788/1 and NE/T010657/1 (SARDINE). DJ was supported by a UK Research and Innovation Future Leaders Fellowship (MR/T020822/1). AM was supported by NERC Grant NE/T01069X/1 (SARDINE) and the BIOPOLE National Capability Multicentre Round 2 funding from the NERC Grant NE/W004933/1.

References

Armour, K. C., Marshall, J., Scott, J. R., Donohoe, A., & Newsom, E. R. (2016). Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nature Geoscience*, *9*(7), 549–554. https://doi.org/10.1038/ngeo2731

Boland, E. J., Jones, D. C., Meijers, A. J., Forget, G., & Josey, S. A. (2021). Local and remote influences on the heat content of Southern Ocean mode water formation regions. *Journal of Geophysical Research: Oceans*, *126*(4), e2020JC016585. https://doi.org/10.1029/2020jc016585

Cerovečki, I., & Meijers, A. J. (2021). Strong quasi-stationary wintertime atmospheric surface pressure anomalies drive a dipole pattern in the Subantarctic Mode Water formation. *Journal of Climate*, *34*(17), 6989–7004.

Czeschel, L., Marshall, D. P., & Johnson, H. L. (2010). Oscillatory sensitivity of Atlantic overturning to high-latitude forcing. *Geophysical Research Letters*, *37*(10), L10601. https://doi.org/10.1029/2010gl043177

ECCO Consortium, Fukumori, I., Wang, O., Fenty, I., Forget, G., Heimbach, P., et al. (2020). Synopsis of the ECCO central production global ocean and sea-ice state estimate (version 4 release 4). *Zenodo*.

Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2015). ECCO version 4: An integrated framework for non-linear inverse modeling and global ocean state estimation. *Geoscientific Model Development*, *8*(10), 3071–3104. https://doi.org/10.5194/gmd-8-3071-2015

Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2016). ECCO version 4: Second release.

Forget, G., Ferreira, D., & Liang, X. (2015). On the observability of turbulent transport rates by Argo: Supporting evidence from an inversion experiment. *Ocean Science*, *11*(5), 839–853. https://doi.org/10.5194/os-11-839-2015

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., & Winton, M. (2015). Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *Journal of Climate*, *28*(2), 862–886. https://doi.org/10.1175/jcli-d-14-00117.1

Fukumori, I., Menemenlis, D., & Lee, T. (2007). A near-uniform basin-wide sea level fluctuation of the Mediterranean Sea. *Journal of Physical Oceanography*, *37*(2), 338–358. https://doi.org/10.1175/jpo3016.1

Gao, L., Rintoul, S. R., & Yu, W. (2018). Recent wind-driven change in Subantarctic Mode Water and its impact on ocean heat storage. *Nature Climate Change*, *8*(1), 58–63. https://doi.org/10.1038/s41558-017-0022-8

Gruber, N., Landschützer, P., & Lovenduski, N. S. (2019). The variable Southern Ocean carbon sink. *Annual Review of Marine Science*, *11*(1), 159–186. https://doi.org/10.1146/annurev-marine-121916-063407

Haumann, F. A., Gruber, N., Münnich, M., Frenger, I., & Kern, S. (2016). Sea-ice transport driving Southern Ocean salinity and its recent trends. *Nature*, *537*(7618), 89–92. https://doi.org/10.1038/nature19101

Hendry, K. R., & Brzezinski, M. A. (2014). Using silicon isotopes to understand the role of the Southern Ocean in modern and ancient biogeochemistry and climate. *Quaternary Science Reviews*, *89*, 13–26. https://doi.org/10.1016/j.quascirev.2014.01.019

Herraiz-Borreguero, L., & Rintoul, S. R. (2011). Subantarctic mode water: Distribution and circulation. *Ocean Dynamics*, *61*(1), 103–126. https://doi.org/10.1007/s10236-010-0352-9

Jones, D. C., Boland, E., Meijers, A. J., Forget, G., Josey, S., Sallée, J.-B., & Shuckburgh, E. (2020). The sensitivity of Southeast Pacific heat distribution to local and remote changes in ocean properties. *Journal of Physical Oceanography*, *50*(3), 773–790. https://doi.org/10.1175/jpo-d-19-0155.1

Jones, D. C., Boland, E., Meijers, A. J., Forget, G., Josey, S. A., Sallee, J.-B., & Shuckburgh, E. (2019). Heat distribution in the Southeast Pacific is only weakly sensitive to high–latitude heat flux and wind stress. *Journal of Geophysical Research: Oceans*, *124*(12), 8647–8666. https://doi.org/10.1029/2019jc015460

Jones, D. C., Forget, G., Sinha, B., Josey, S. A., Boland, E. J., Meijers, A. J., & Shuckburgh, E. (2018). Local and remote influences on the heat content of the Labrador sea: An adjoint sensitivity study. *Journal of Geophysical Research: Oceans*, *123*(4), 2646–2667. https://doi.org/10.1002/2018jc013774

Khatri, H., Williams, R. G., Woollings, T., & Smith, D. M. (2022). Fast and slow subpolar ocean responses to the North Atlantic Oscillation: Thermal and dynamical changes. *Geophysical Research Letters*, *49*(24), e2022GL101480. https://doi.org/10.1029/2022gl101480

Kostov, Y., Johnson, H. L., & Marshall, D. P. (2019). AMOC sensitivity to surface buoyancy fluxes: The role of air-sea feedback mechanisms. *Climate Dynamics*, *53*(7–8), 4521–4537. https://doi.org/10.1007/s00382-019-04802-4

Kwon, E. Y., Downes, S. M., Sarmiento, J. L., Farneti, R., & Deutsch, C. (2013). Role of the seasonal cycle in the subduction rates of upper–Southern Ocean waters. *Journal of Physical Oceanography*, *43*(6), 1096–1113. https://doi.org/10.1175/jpo-d-12-060.1

Large, W. G., & Yeager, S. G. (2004). *Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies*. National Center for Atmospheric Research Boulder.

Marotzke, J., Giering, R., Zhang, K. Q., Stammer, D., Hill, C., & Lee, T. (1999). Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity. *Journal of Geophysical Research*, *104*(C12), 29529–29547. https://doi.org/10.1029/1999jc900236

Marshall, J., & Speer, K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience*, *5*(3), 171–180. https://doi.org/10.1038/ngeo1391

McCartney, M. (1977). Subantarctic mode water. A voyage of discovery: George deacon 70th anniversary volume, MV Angel, Ed. *Deep-Sea Research Supplement*, 103–119.

Meijers, A., Cerovečki, I., King, B. A., & Tamsitt, V. (2019). A see-saw in Pacific subantarctic mode water formation driven by atmospheric modes. *Geophysical Research Letters*, *46*(22), 13152–13160. https://doi.org/10.1029/2019gl085280

Morrison, A. K., Waugh, D. W., Hogg, A. M., Jones, D. C., & Abernathey, R. P. (2022). Ventilation of the Southern Ocean pycnocline. *Annual Review of Marine Science*, *14*(1), 405–430. https://doi.org/10.1146/annurev-marine-010419-011012

Ogle, S., Tamsitt, V., Josey, S., Gille, S., Cerovečki, I., Talley, L., & Weller, R. (2018). Episodic Southern Ocean heat loss and its mixed layer impacts revealed by the farthest south multiyear surface flux mooring. *Geophysical Research Letters*, *45*(10), 5002–5010. https://doi.org/10.1029/2017gl076909

Pillar, H. R., Heimbach, P., Johnson, H. L., & Marshall, D. P. (2016). Dynamical attribution of recent variability in Atlantic overturning. *Journal of Climate*, *29*(9), 3339–3352. https://doi.org/10.1175/jcli-d-15-0727.1

Pimm, C. (2023). Data for 'Surface heat fluxes drive a two-phase response in Southern Ocean mode water stratification' [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.8160115

Sallée, J.-B., Speer, K., Rintoul, S., & Wijffels, S. (2010). Southern Ocean thermocline ventilation. *Journal of Physical Oceanography*, *40*(3), 509–529. https://doi.org/10.1175/2009jpo4291.1

Smith, T., & Heimbach, P. (2019). Atmospheric origins of variability in the South Atlantic meridional overturning circulation. *Journal of Climate*, *32*(5), 1483–1500. https://doi.org/10.1175/jcli-d-18-0311.1

Talley, L., & Raymer, M. (1982). Eighteen degree water variability. *Journal of Marine Research*, *40*, 757–775.

Tamsitt, V., Cerovečki, I., Josey, S. A., Gille, S. T., & Schulz, E. (2020). Mooring observations of air–sea heat fluxes in two Subantarctic Mode Water formation regions. *Journal of Climate*, *33*(7), 2757–2777. https://doi.org/10.1175/jcli-d-19-0653.1

Verdy, A., Mazloff, M. R., Cornuelle, B. D., & Kim, S. Y. (2014). Wind-driven sea level variability on the California coast: An adjoint sensitivity analysis. *Journal of Physical Oceanography*, *44*(1), 297–318. https://doi.org/10.1175/jpo-d-13-018.1

Williams, R. G. (1988). Modification of ocean eddies by air–sea interaction. *Journal of Geophysical Research*, *93*(C12), 15523–15533. https://doi.org/10.1029/jc093ic12p15523

Williams, R. G., Ceppi, P., Roussenov, V., Katavouta, A., & Meijers, A. J. (2023). The role of the Southern Ocean in the global climate response to carbon emissions. *Philosophical Transactions of the Royal Society A*, *381*(2249), 20220062. https://doi.org/10.1098/rsta.2022.0062

Williams, R. G., & Meijers, A. (2019). In J. K. Cochran, H. J. Bokuniewicz, & P. L. Yager (Eds.), *Ocean subduction. Encyclopedia of ocean sciences* (3rd ed., pp. 141–157). Academic Press.