Abstract

A holistic review is given of the Southern Ocean dynamic system, in the context of the crucial role it plays in the global climate and the profound changes it is experiencing. The review focuses on connections between different components of the Southern Ocean dynamic system, drawing together contemporary perspectives from different research communities, with the objective of ‘closing loops’ in our understanding of the complex network of feedbacks in the overall system. For the purposes of this review, the Southern Ocean dynamic system is divided into four main components: large-scale circulation; cryosphere; turbulence; and gravity waves. Overviews are given of the key dynamical phenomena for each component, before describing the linkages.
between the components. The reviews are complemented by an overview of observed Southern Ocean trends and future climate projections. Priority research areas required to improve our understanding of the Southern Ocean system are identified.

Abstract content goes here
Closing the loops on Southern Ocean dynamics: From the circumpolar current to ice shelves and from bottom mixing to surface waves

Key Points:

• Contemporary perspectives on the different components of the Southern Ocean dynamic system from distinct research communities are reviewed
• Key connections between different components of Southern Ocean dynamics are highlighted
• Cross-cutting priorities for future Southern Ocean physical science are identified

Corresponding author: Luke G. Bennetts, luke.bennetts@adelaide.edu.au
Abstract
A holistic review is given of the Southern Ocean dynamic system, in the context of the crucial role it plays in the global climate and the profound changes it is experiencing. The review focuses on connections between different components of the Southern Ocean dynamic system, drawing together contemporary perspectives from different research communities, with the objective of ‘closing loops’ in our understanding of the complex network of feedbacks in the overall system. For the purposes of this review, the Southern Ocean dynamic system is divided into four main components: large-scale circulation; cryosphere; turbulence; and gravity waves. Overviews are given of the key dynamical phenomena for each component, before describing the linkages between the components. The reviews are complemented by an overview of observed Southern Ocean trends and future climate projections. Priority research areas required to improve our understanding of the Southern Ocean system are identified.

Plain Language Summary
The United Nations has identified 2021–2030 as the Decade of Ocean Science, with a goal to improve predictions of ocean and climate change. Improved understanding of the Southern Ocean is crucial to this effort, as it is the central hub of the global ocean. The Southern Ocean is the formation site for the dense water that fills the deep ocean, sequesters the majority of anthropogenic heat and carbon, and controls the flux of heat to Antarctica. The large-scale circulation of the Southern Ocean is strongly influenced by interactions with sea ice and ice shelves, and is mediated by smaller scale processes, including eddies, waves and mixing. The complex interplay between these dynamic processes remains poorly understood, limiting our ability to understand, model and predict changes to the Southern Ocean, global climate and sea level. This article provides a holistic review of Southern Ocean processes, connecting the smallest scales of ocean mixing to the global circulation and climate. It seeks to develop a common language and knowledge-base across the Southern Ocean physical science community to facilitate knowledge-sharing and collaboration, with the aim of closing the loops on our understanding of one of the world’s most dynamic regions.
1 Introduction

The Southern Ocean is a harsh, dynamic and remote environment, which has profound influence over Earth’s present and future climates. It is home to the global ocean’s strongest winds, coldest ocean surface temperatures, largest ice shelves, biggest ocean currents, and most extreme surface waves. The Southern Ocean connects other major global oceans to the north, acting as a central hub where waters from the Atlantic, Pacific and Indian basins converge and mix. As such, it regulates the uptake of heat and carbon at a global scale. To the south, the unique dynamics of the Southern Ocean control the flux of heat to Antarctica’s fringes, thus stabilising the Antarctic Ice Sheet (which holds the volumetric equivalent of about 60 m in global mean sea level; Fretwell et al., 2013; Morlighem et al., 2020). However, the Southern Ocean is experiencing profound, large-scale changes, many at unprecedented and accelerating rates. These include the lowest ever recorded sea ice minima in the past two Austral summers (NISDC, 2023), rapid melting of the West Antarctic Ice Sheet (Paolo et al., 2015), and the warming and freshening of the abyssal waters formed in the Southern Ocean ( Purkey & Johnson, 2013).

The observed large-scale changes in the Southern Ocean climate assimilate a rich spectrum of dynamics, spanning thousand-kilometre scale ocean currents, hundred-kilometre scale polynyas, ten-kilometre wide eddies, kilometre-scale convection, hundred-metre scale surface waves, metre-scale pancake sea ice and millimetre-scale turbulent mixing. The network of linkages and feedbacks between the different components of the Southern Ocean dynamic system creates challenges in understanding and predicting this vitally important region and its role in global climate and ecosystems. The objective of this review is to ‘close loops’ on the Southern Ocean dynamic system by drawing together contemporary perspectives on the different components of the system from different research communities. The focus is on a holistic physical understanding of the Southern Ocean, rather than associated aspects of atmospheric dynamics, land-based ice, and dynamical interactions with biogeochemistry. Instead, the reader is directed to reviews by Noble et al. (2020) for Antarctic Ice Sheet dynamics and Henley et al. (2020) for Southern Ocean biogeochemistry. In addition, there exist a number of reviews into different aspects of atmospheric dynamics and air–sea coupling, including the Southern Annular Mode (Fogt & Marshall, 2020), Southern Ocean precipitation (Siems et al., 2022) and air–sea–ice exchanges (S. Swart et al., 2019).

There are several definitions of the Southern Ocean extent; we take a dynamical perspective and consider the Southern Ocean system to be bounded by the northern most extent of the Antarctic Circumpolar Current, and that its southern boundary includes the ice shelves and sub-ice shelf cavities fringing the Antarctic continent. In the vertical direction, we consider dynamics stretching from the ocean surface, which is occupied by surface gravity waves and sea ice cover, to the ocean bottom, which is a key region for the generation of internal waves and subsequent mixing. We divide the Southern Ocean dynamic system into four main components: large-scale circulation; cryosphere; turbulence; and gravity waves. Large-scale circulation incorporates the Antarctic Circumpolar Current, Antarctic Slope Current, sub-polar gyres, and the meridional overturning circulation. The cryosphere includes sea ice and ice shelves, as well as dynamic phenomena in the sub-shelf cavities. We define turbulence as chaotic dynamics spanning from mesoscale eddies and polynya convection at the largest end, down to millimetre-scale diapycnal mixing. Gravity waves includes surface waves, internal waves and tides. Fig. 1 gives a spatio-temporal perspective of phenomena reviewed, which shows the broad range of scales covered. We focus the review on the connected nature of interactions between the different phenomena.

We structure the review around the four main dynamical components identified above, commencing with large-scale circulation (§ 2) to provide a global perspective on the Southern Ocean dynamical environment, followed by cryosphere (§ 3), turbulence (§ 4) and gravity waves (§ 5). In each section, we give an overview of the fundamental physics of the dynamical component being considered, before describing the linkages between these components. In prioritising these linkages, we focus on the most impactful, those in areas of growing
Figure 1. Diagram showing a spatio-temporal perspective on key dynamic phenomena reviewed, where colours indicate association to the four key dynamical components.

research activity, and those where significant outstanding questions remain. We typically describe the linkages in the section corresponding to the component that is being impacted, thereby minimising repetition. The sections dedicated to the four dynamical components are followed by an overview of relevant Southern Ocean climate trends and future climate projections (§6). We close the review with a summary of our present understanding of Southern Ocean dynamics and by identifying cross-cutting priorities for future Southern Ocean physical science (§7).
2 Large-scale circulation

Figure 2. Schematic of the Southern Ocean’s large-scale circulation, where the ocean colours indicate the density, ranging from lighter (dark orange) to denser (dark blue) waters, and isopycnal contours are the interfaces between the layers. The horizontal gradients in density are correlated with largely geostrophic currents, including the Antarctic Circumpolar Current (ACC) and Antarctic Slope Current (ASC), above the shelf slope/break. Antarctic Bottom Water is generated by convection and brine rejection on the continental shelf, and flows down into the abyssal ocean. Warmer Circumpolar Deep Water (CDW) is upwelled in the mid-depths and plays a key role in the melt rate of ice shelves. These processes collectively form the Southern Ocean component of the upper and abyssal overturning cells, as indicated by the dashed lines. Farther to the north, at the density fronts of the ACC, are the formation sites of northward flowing mode and intermediate waters. The topography, isopycnals, and ice shelf profile on the southern side of the schematic are from observations in the Ross Sea, although they are artificially extended to the north to represent a more typical condition for the Antarctic Circumpolar Current. Note that the depth scale is nonuniform.

Large-scale circulation is here interpreted as flows at horizontal scales larger than mesoscale eddies (greater than \( \sim 300 \text{ km} \)). An schematic representation of large-scale circulation in the vertical–latitude plane (Fig. 2) identifies the so-called “meridional overturning circulation”, which includes the upper (clockwise in Fig. 2; §2.4) and abyssal (anticlockwise; §2.5) branches. Figure 3 shows a plan view of the entire Southern Ocean to highlight the horizontal circulation features: the Antarctic Circumpolar Current (§2.1), Antarctic Slope Current (§2.2), and Weddell and Ross gyres (§2.3).

The large-scale circulations are broadly in geostrophic balance, although multi-scale interactions play a fundamental role in their variability and response to forcing. The aim of this section is to offer a perspective on the processes involved in sustaining these circulations and the links that bring them together. In broad terms, they are sustained through multi-scale interactions between mean flows, turbulence, topography, dynamic stresses, isopycnal mixing and buoyancy fluxes. The understanding of how the exchange of tracers, momentum, and vorticity connect the different components of the large-scale Southern Ocean circulation together is rapidly evolving. The reader is referred to previous reviews on Southern Ocean circulation for more details on specific processes. In particular, Rintoul and Naveira Gara-
Figure 3. A plan view of the large-scale circulation of the Southern Ocean. The key features are the braided network of eddies and jets circumnavigating the continent that comprise the eastward-flowing Antarctic Circumpolar Current (for which the northern and southern limits, or fronts, are represented by black contours), the Weddell and Ross Gyres (blue dotted lines), and the westward-flowing Antarctic Slope/Coastal Current nearer the continent (brown line). The Slope Current exists everywhere except along the western side of the Antarctic Peninsula, where the Circumpolar Current flows very close to the shelf slope. The background image shows a typical snapshot of daily mean surface flow speed from the ACCESS-OM2-01 global ocean model (Kiss et al., 2020).

bato (2013) provide a detailed discussion of the Southern Ocean’s role in the global ocean circulation and climate, while A. F. Thompson et al. (2018), Vernet et al. (2019) and Dotto et al. (2018) provide detailed reviews of the Antarctic Slope Current, Weddell and Ross Gyres, respectively.

2.1 Antarctic Circumpolar Current

The Antarctic Circumpolar Current is the largest ocean current in the world. It encircles Antarctica, extending from the surface to ~2km deep, connecting the Atlantic, Pacific and Indian ocean basins, and forming the hub of the global ocean circulation (e.g., see Fig. 1 of Meredith, 2022). The sloping density surfaces (isopycnals) associated with the Antarctic Circumpolar Current provide a connection between the ocean surface and the abyss. They allow fluid from the deep ocean to upwell without changing its density, which is a crucial component of the global overturning circulation (§§2.4–2.5). The regions of
sharpest meridional density gradient at the surface are described as the (density) fronts of
the Antarctic Circumpolar Current, with the northern- and southern-most fronts (Fig. 3)
enclosing the region of strongest current speed.

Despite its importance to the climate system, the dynamics controlling the Antarctic
Circumpolar Current are not entirely understood. The geometry of the current is unique;
unlike other ocean currents, there are no continents blocking its quasi-zonal flow around
the globe. This unique configuration means that the dynamics of the Antarctic Circum-
polar Current cannot be explained using the classical geophysical fluid dynamics theories
that govern gyres, although some have tried to apply these concepts, such as the Sver-
drup balance, to the Antarctic Circumpolar Current (e.g., Stommel, 1957; Webb, 1993;
C. W. Hughes, 1997). The integrated momentum balance of the Antarctic Circumpolar
Current is extremely simple: wind stress at the surface is predominantly balanced by topo-
graphic form stress at the bottom (Masich et al., 2015a), as originally proposed by Munk
and Palmén (1951). Hiding in this simplicity is a wealth of surprises and contradictions. For
example, despite the wind stress being the dominant source of momentum for the Antarctic
Circumpolar Current, changing the wind has almost no effect on the total zonal baroclinic
transport (Straub, 1993; Hallberg & Gnanadesikan, 2001; Tansley & Marshall, 2001; Mun-
day et al., 2013; Constantinou & Hogg, 2019), and increasing the bottom drag increases
the total zonal transport (D. P. Marshall et al., 2017; Constantinou, 2018). Moreover, although
mesoscale turbulence is believed to play a crucial role in fluxing momentum downwards from
the surface to be dissipated at depth, the momentum budget adjusts to wind changes within
a month (Ward & Hogg, 2011; Masich et al., 2015b), while the response of the mesoscale
turbulence is much slower, taking months to years to adjust (Meredith & Hogg, 2006a; Sinha
& Abernathey, 2016; Hogg et al., 2022). Modelling results also suggest that the Antarctic
Circumpolar Current responds in different ways to specific spatial patterns of wind stresses,
such as those associated with the interplay of the different phases of the Southern Annular
Mode and El Nino-Southern Oscillation (Langlais et al., 2015).

The Antarctic Circumpolar Current is often referred to as if it were a single monolithic
current. This interpretation is based on a time-mean field view of the Antarctic Circum-
polar Current. But the instantaneous current is better described as a complex network of
interconnected jets and eddies (§ 4.1). This smaller-scale structure supports a plethora
of multiscale interactions: eddy-jet interactions shorten eddy lifetimes (R. Liu et al., 2022);
jet-topography interactions can lead to rapid changes in ocean ventilation (Klocker, 2018);
and a unique set of interactions occur where the eastward flowing Antarctic Circumpolar
Current in the Southern Ocean is fast enough to arrest westward propagating Rossby waves
(Klocker & Marshall, 2014). Downstream of large bathymetric features the time-mean
flow field exhibits standing meanders, thought to be the result of arrested Rossby waves
(A. F. Thompson & Naveira Garabato, 2014). These arrested Rossby waves also affect the
stability of the current, allowing instabilities to grow when the wave speed matches or ex-
ceeds the flow speed and is oriented in the opposing direction, i.e., the wave is travelling
upstream (Stanley et al., 2020). These standing meander regions are also highly energetic,
with enhanced cross-frontal exchange (A. F. Thompson & Sallée, 2012), eddy heat flux
(Foppert et al., 2017) and upwelling (Tamsitt et al., 2017). The Antarctic Circumpolar
Current flows along standing meanders, whose curved paths lead to horizontal divergence
and vortex stretching that couples the upper and lower water column, modifying deep cur-
rents and cross-frontal exchange in patterns locked to the phase of the meander (Meijer
et al., 2022).

2.2 Antarctic Slope Current

The steep gradient of the Antarctic continental shelf slope imposes a strong geometric
constraint on cross-slope flow as it invokes a large potential vorticity gradient. Consequently,
ocean flows in this region are (to first order) oriented in an along-slope direction, known as
the Antarctic Slope Current. The Antarctic Slope Front is the associated front and is man-
ifested by a large cross-slope density gradient. The Antarctic Slope Current is strongest in East Antarctica and exists everywhere along the shelf slope except for the western Antarctic Peninsula, where the Antarctic Slope Current is replaced by the southernmost edge of the Antarctic Circumpolar Current (Mathiot et al., 2011; Armitage et al., 2018; A. L. Stewart et al., 2019; Pauthenet et al., 2021; Huneke et al., 2022). The Antarctic Slope Current advects tracers such as heat, salt and nutrients around the continent, and the exchange of very distinct water masses across the current is pivotal for the climate system. Onshore transport of relatively warm Circumpolar Deep Water impacts ice shelf basal melt rates and pre-conditions Dense Shelf Water characteristics (§3.1.1). Offshore transport of Dense Shelf Water ventilates the deep ocean by supplying water for Antarctic Bottom Water, feeding into the deep overturning circulation (§2.5).

The almost circumpolar extent of the Antarctic Slope Current and its importance for the climate system was established in the 20th century (Gill, 1973; Whitworth et al., 1998; Jacobs, 1991). The advancement of numerical ocean model capabilities over the past decade, as well as increased efforts to collect observations (ship-based, moorings/fixed, animal-borne, autonomous vehicles), has led to a rapid expansion of the research field and an improved understanding of the Antarctic Slope Current dynamics. This section builds on the review by A. F. Thompson et al. (2018), highlighting recent developments, with a focus on interaction with the cryosphere (§3), as well as touching on the mesoscale eddy (§4.1) and tide (§5.2) influences.

The Antarctic Slope Current is driven primarily by winds and buoyancy forcing from both the atmosphere and meltwater. At leading order, easterly winds around Antarctica are oriented in an along-slope direction (Hazel & Stewart, 2019), driving onshore Ekman transport, creating a cross-slope density gradient, and thereby driving an along-slope current in thermal wind balance. The momentum transfer from the atmosphere to the ocean occurs via the sea ice that covers the continental shelf for most of the year. Recent high resolution model simulations indicate that the surface stress over the continental shelf vanishes in the presence of sea ice (A. L. Stewart et al., 2019; Si et al., 2021), and the sea ice distributes the momentum input provided by the wind away from the continental slope. In addition to winds, buoyancy fluxes from the sea ice, ice shelves and atmosphere help sustain the cross-slope pressure gradients that support the Antarctic Slope Current. Freshwater forcing from ice shelf melting plays a particularly important role (Fahrbach et al., 1992; Moffat et al., 2008), with new observations suggesting that glacial melt is especially important for the generation of the Antarctic Slope Current in the Amundsen Sea (A. F. Thompson et al., 2020). This mechanism is supported by model simulations with amplified freshwater forcing (to represent basal melting of ice shelves), which show an increased cross-slope density gradient and enhanced Antarctic Slope Current (Naughten et al., 2018; Moorman et al., 2020; Beadling et al., 2022). The Antarctic Slope Current is reinforced by tides through a process called tidal rectification (§5.2.3 A. L. Stewart et al., 2019; Si et al., 2021).

The state of the Antarctic Slope Current is closely related to Dense Shelf Water export, which occurs downstream of the Ross Sea, Adelie Land, Prydz Bay, and the Weddell Sea (A. F. Thompson et al., 2018). The presence of dense water lifts the isopycnals at depth, connecting the shelf with the offshore ocean and creating a pathway for eddy-driven cross-slope heat exchange (A. L. Stewart & Thompson, 2015). Further, the Dense Shelf Water descending the continental shelf gives rise to a bottom-intensified Antarctic Slope Current flow in these locations, unlike other regions where it is surface intensified (e.g., Heywood et al., 1998; Huneke et al., 2022).

### 2.3 Weddell and Ross Gyres

The Weddell and Ross Gyres are dominant features of the lateral circulation of the Southern Ocean, located south of the Antarctic Circumpolar Current and north of the Antarctic continental shelf (Fig. 4). These gyres play a mediating role in the exchange
of waters between the relatively warm waters within the Antarctic Circumpolar Current and the cold continental shelf. Both of these gyres are also located adjacent to one of the formation sites of Dense Shelf Water around Antarctica (Purkey et al., 2018; Meredith, 2013). Thus, the properties of exported Antarctic Bottom Water can also be influenced by gyre circulation (Bai et al., 2022; Meredith et al., 2014), as can be the properties of the source waters that participate in Dense Shelf Water production (Narayanan et al., 2019; Foster & Carmack, 1976). Therefore, there is a connection between the Ross and Weddell Gyre circulation and processes relevant to global climate, such as ocean heat and carbon uptake (MacGilchrist et al., 2019; P. J. Brown et al., 2015). The circulation of the gyres has also been found to influence polynya formation (Zhou et al., 2022; Cheon & Gordon, 2019; Cheon et al., 2018), sea ice variability (Morioka & Behera, 2021; Neme et al., 2021) and iceberg drift (Barbat et al., 2021; Bouhier et al., 2018).

The average climatological wind field makes the gyres a region of divergent Ekman transport, fostering a vertical structure characterised by isopycnals sloping upwards towards the centre of the gyre, with local upwelling and mixing of subsurface Circumpolar Deep Water (Jullion et al., 2014). Circumpolar Deep Water is able to enter the gyres through permeable eastern boundaries, where there is no topographic constraint to their circulation (Bebieva & Speer, 2021; Roach & Speer, 2019; Donnelly et al., 2017; Ryan et al., 2016; Cisewski et al., 2011; Fährbach et al., 2011; A. H. Orsi & Wiedermühl, 2009). Because of this lack of topographic constraint, the eastern extent of the gyres is highly variable and uncertain (Wilson et al., 2022; Vernet et al., 2019; Dotto et al., 2018; Roach & Speer, 2019), with eddies and high frequency variability associated with topographic discontinuities playing an important role in the exchange of waters (Bebieva & Speer, 2021; Roach & Speer, 2019; Donnelly et al., 2017; Ryan et al., 2016). Within the gyres, Circumpolar Deep Water is shielded from interaction with the atmosphere by a shallow layer of colder and fresher water that builds up during winter and erodes during summer. Upwelling and entrainment via diapycnal mixing of warm and salty Circumpolar Deep Water into the surface layer contributes to sea ice melt (Bebieva & Speer, 2021; Wilson et al., 2019) and polynya formation (§ 4.2.3; Campbell et al., 2019).

As Dense Shelf Water cascades down the continental shelf it enters the gyres and undergoes further transformation as it becomes entrained with ambient waters to produce Antarctic Bottom Water (Akhoudas et al., 2021; Gordon et al., 2009). In the Weddell Gyre, it has been suggested that the properties and rates of export of Antarctic Bottom Water across gyre boundaries are dependent on the gyre’s horizontal circulation due to two different mechanisms. From a baroclinic perspective, an acceleration of the gyre induces an increase in the isopycnal tilt at its northern boundary, effectively trapping the densest varieties of bottom waters that are not able to overflow through the shallow passages (Gordon et al., 2009; Meredith et al., 2008). From a barotropic perspective, an acceleration of the gyre induces changes in the strength of the deep boundary current near the outflow locations (Meredith et al., 2011). The Ross Gyre is more sparsely observed, but there is evidence that the gyre’s circulation modulates the salinity of the Dense Shelf Water formed at the western Ross Sea (Guo et al., 2020) and induces changes in the properties and volume of Antarctic Bottom Water in the Amundsen and Bellingshausen Seas (Bai et al., 2022). Warm intrusions of Circumpolar Deep Water onto the Amundsen and Bellingshausen shelf are also related to the Ross Gyre’s strength (Nakayama et al., 2018). In addition to dense waters, meltwater coming from the ice shelves in the Ross and Weddell Seas is also partly distributed within the gyres’ circulation (Kusahara & Hasumi, 2014).

There are few studies addressing the variability of the Ross and Weddell Gyres across different time scales in connection to possible forcing mechanisms. Satellite-based studies have found links between the gyre’s sea surface height and wind stress curl (Auger, Sallée et al., 2022; Armitage et al., 2018). However, the extensive sea ice coverage in the region means that the transfer of momentum from the wind to the ocean surface is modulated by sea ice, which has to be taken into account when considering surface stresses (Neme
Figure 4. Schematics for circulation of the (a) Weddell and (b) Ross Gyres. Dense Shelf Water (DSW) formation regions are highlighted in purple (LCIS: Larsen C Ice Shelf, FRIS: Filchner-Ronne Ice Shelf, RIS: Ross Ice Shelf) with Dense Shelf Water pathways indicated with purple arrows. Red arrows indicate the Antarctic Circumpolar Current (ACC), orange arrows indicate pathways of Circumpolar Deep Water (CDW) with orange circles marking eddy-rich regions, and yellow arrows mark the circulation of modified Circumpolar Deep Water within the gyres. Blue arrows indicate Antarctic Bottom Water (AABW) pathways.

et al., 2021; A. Naveira Garabato et al., 2019; Dotto et al., 2018). By including sea ice in the total stress over the ocean’s surface, the correlation with sea surface height breaks down (Auger, Sallée, et al., 2022), as it does with the gyre’s strength on both seasonal and interannual timescales (Neme et al., 2021). There are different processes within the gyres that could be playing a role in their variability, thus obscuring a direct relation with surface stress, such as variability of water mass exchange across gyre boundaries or variability of dense water formation. Führbach et al. (2011) suggest that the northern and southern limbs of the Weddell Gyre can vary independently due to asymmetric wind forcing. Moreover, there are studies suggesting that ocean gyres can develop in response to surface buoyancy
fluxes (Hogg & Gayen, 2020; Bhagtani et al., 2023). In support of this hypothesis, the Ross and Weddell Gyre strengths in climate models have been found to be correlated with the upper layer meridional density gradients, whilst being largely independent of wind stress curl (Z. Wang & Meredith, 2008).

2.4 Upper overturning circulation

The upper overturning circulation of the Southern Ocean consists of southward upwelling flow along steeply tilted isopycnals in the mid-depths and a return northward flow of lighter waters (called “mode” or “intermediate” waters, due to their density being intermediate between abyssal and surface waters) in the upper ocean (Fig. 2). The lifting (or “upwelling”) of deep waters to the surface and subsequent subduction north of the Antarctic Circumpolar Current has a large impact on global climate by enabling rapid exchange of heat and carbon between the atmosphere and interior ocean (Morrison et al., 2015). A large fraction of the global ocean uptake of anthropogenic heat (∼70%) and carbon (∼40%) has occurred in the Southern Ocean, due to the constant replenishment of surface waters with colder and carbon-depleted water from below (Frölicher et al., 2015; Zanna et al., 2019; Khatiwala et al., 2009).

Buoyancy fluxes (i.e., the combined effect of sensible, latent, radiative and freshwater fluxes) and wind stresses at the ocean surface have a strong control over the strength and structure of the upper overturning circulation. The westerly winds drive Ekman upwelling along isopycnals that outcrop south of the maximum wind stress (∼55°S) and downwelling along isopycnals to the north (J. R. Toggweiler & Samuels, 1993; Speer et al., 2000; J. Marshall & Speer, 2012). The overturning transport increases with increasing wind stress (e.g., Viebahn & Eden, 2010; Bishop et al., 2016), though the sensitivity is less than the Ekman transport response due to the additional impact of buoyancy forcing and eddies on the dynamics (e.g., Abernathey et al., 2011). Buoyancy input, predominantly from sea ice melt and precipitation, transforms the dense upwelled waters into lighter, northward flowing waters at the surface (Abernathey et al., 2016). Surface buoyancy forcing also plays a critical role in the formation of mode waters on the northern edge of the Antarctic Circumpolar Current in the Indian and Pacific sectors (Wong, 2005; Sloyan & Rintoul, 2001; Sallée et al., 2010). In particular, surface cooling and evaporation drive strong wintertime convection, forming Subantarctic Mode Water (Hanawa & Talley, 2001; Abernathey et al., 2016). The shoaling of the deep mixed layers during spring then results in a net subduction of waters from the mixed layer to beneath the permanent pycnocline (Z. Li et al., 2022; Morrison et al., 2022).

Eddies (§ 4.1) play a critical role in the upwelling branch of the overturning circulation. Southward flow in the mid-depths of the Southern Ocean is dominated by eddy transport along isopycnals, due to the lack of land barriers required for zonal mean geostrophic flows in the meridional direction (J. Marshall & Speer, 2012). The generation of eddy kinetic energy through baroclinic instability extracts available potential energy from the sloping isopycnals. This energy conversion results in a flattening of the isopycnals and, therefore, a net southward (and upwards) transport in the upper and mid-depth ocean (Morrison et al., 2015). The southward flow has a highly heterogeneous spatial distribution around the Southern Ocean, with southward transport collocated with baroclinic eddy generation downstream (eastward) of topographic hotspots (Tamsitt et al., 2017; Barthel et al., 2022; Yung et al., 2022). The hotspots of eddy generation and southward transport are located ∼100 km upstream (westward) of the eddy kinetic energy hotspots (Foppert et al., 2017; Yung et al., 2022). Eddies impact the dynamics of the upper overturning circulation by limiting the sensitivity of the overturning transport to changing wind stress (e.g., Hallberg & Gnanadesikan, 2006; Viebahn & Eden, 2010; Gent, 2016), and by influencing the formation rate of mode and intermediate waters (Sallée et al., 2010; Z. Li et al., 2022). The eddy response in limiting the overturning circulation sensitivity to wind changes is known as “eddy compensation” (Fig. 5). Eddy compensation occurs because the southward eddy
transport extends all the way into the surface layers, directly opposing the northward Ekman transport in the upper ocean (J. Marshall & Radko, 2003). Following an increase in westerly wind stress, the northward Ekman transport and the southward eddy transport in the surface layers both increase, resulting in a reduced sensitivity of the overturning to wind stress compared to a hypothetical situation with no change in eddy activity. However, the overturning transport (i.e., the maximum value of the zonal-mean overturning streamfunction in latitude-depth coordinates) still increases with increasing wind stress, because much of the southward eddy transport occurs below the Ekman layer and does not play a role in the compensation (Morrison & Hogg, 2013).

![Figure 5](image-url)

**Figure 5.** Simulated eddy compensation of the upper overturning circulation. With no eddy compensation (by resolved or parameterised eddies), the overturning would linearly increase with the magnitude of the westerly wind stress following the surface Ekman transport (black dashed line). As model resolution is increased such that mesoscale eddies become fully resolved (red line), the sensitivity of the overturning circulation to wind stress decreases, but remains non-zero. Figure reproduced from Morrison and Hogg (2013).

Isopycnal mixing (§ 4.3) and the resultant densification (known as cabbeling) play a first-order role in the formation of Antarctic Intermediate Water (Nycander et al., 2015; Groeskamp et al., 2016; Z. Li et al., 2022), which is a mode water that forms part of the northward return limb of the upper overturning circulation (Fig. 2). Mesoscale eddies stir the strong along-isopycnal temperature and salinity gradients in the Southern Ocean. Due to the nonlinear dependence of water density on temperature, the resultant mixing of surface and interior water parcels with different temperature and salinity, but equal density, produces a water parcel that is denser than the two original water parcels. The key role of cabbeling is illustrated by comparing numerical model simulations with linear and nonlinear equations of state; the former cannot reproduce the salinity minimum associated with Antarctic Intermediate Water (Fig. 6; Nycander et al., 2015).

### 2.5 Abyssal overturning circulation

The Southern Ocean abyssal overturning circulation is considered, in a zonally integrated sense, to consist of two compensating flows: (i) a poleward flow of Circumpolar Deep
Figure 6. Impact of the nonlinear equation of state (i.e., the equation describing the dependence of the density of seawater on temperature, salinity and pressure) on simulated Antarctic Intermediate Water formation. (a–b) Latitudinal transects along $23.5^\circ$W of salinity (colour, with contours labelled in black) and potential density (white labelled contours) in the South Atlantic. The simulation shown in (a) uses a full non-linear equation of state, while (b) uses a linear equation of state. Antarctic Intermediate Water (blue to green freshwater pathway shown in a) forms through isopycnal mixing leading to cabbeling and is only able to form in the model configuration using a nonlinear equation of state. Figure reproduced from Nycander et al. (2015).

Water; and (ii) an equatorward flow of Antarctic Bottom Water (Fig. 2). Circumpolar Deep Water is modified by mixing as it travels poleward to the Antarctic continental shelf, where it is transformed into Dense Shelf Water through surface buoyancy fluxes and brine rejection due to sea ice formation. Dense Shelf Water mixes with and entrains Circumpolar Deep Water as it descends into the abyssal ocean to form Antarctic Bottom Water (A. Orsi et al., 1999). The resulting water mass accounts for 30–40% of the ocean’s total volume, and fills the abyssal depths of the Atlantic, Pacific and Indian Oceans with carbon- and oxygen-rich water (Johnson, 2008). It is estimated that the maximum northward Antarctic Bottom Water transport is about 20–30 Sv near $30^\circ$S (Ganachaud et al., 2000; Lumpkin & Speer, 2007; Talley et al., 2003; Talley, 2008, 2013). The remainder of §2.5 provides an overview of the physical processes that influence the abyssal overturning and the multi-scale links that bring them together.

The cryosphere influences abyssal overturning by modulating Dense Shelf Water formation through three main pathways: ice shelves, sea ice, and major cryospheric events. Ice shelf basal melting sets the location polynyas within areas of landfast ice (Nihashi & Ohshima, 2015), and glacial meltwater has been connected to changes in Antarctic Bottom Water properties (§6.1). Brine rejection during sea ice formation influences the amount of Dense Shelf Water formation, and its salinity and density (e.g., Jacobs, 2004; Iudicone et al., 2008; Abernathey et al., 2016; Silvano et al., 2020). Large cryospheric events, such as the calving of the Mertz Glacier Tongue (Tamura et al., 2012; Shadwick et al., 2013; Aoki et al., 2017; Snow et al., 2018) or the opening of the Weddell Sea polynya (Martinson, 1991;
Akhoudas et al., 2021), reorganise the circulation and stratification and, therefore, alter Dense Shelf Water formation.

The export of Dense Shelf Water occurs predominantly in submerged canyons that cross the continental shelf (Nakayama, Ohshima, et al., 2014). Dense Shelf Water accumulates in these deeper sections of the shelf and eventually spills down the shelf slope, sometimes in short bursts lasting a few days (Foppert et al., 2021). The export of Dense Shelf Water is modulated by tidal mixing, which modifies the water mass properties and helps to bring Circumpolar Deep Water onshore (Muench et al., 2009; Q. Wang et al., 2013; Bowen et al., 2021). Morrison et al. (2020) find that the Circumpolar Deep Water inflow is partly forced by a barotropic pressure gradient set up by the overflowing Dense Shelf Water. Eddies are also a major contributor to Dense Shelf Water and Circumpolar Deep Water transport across the continental slope (A. L. Stewart & Thompson, 2015; Q. Wang et al., 2009; Nakayama, Ohshima, et al., 2014; Nøst et al., 2011). The Dense Shelf Water component of Antarctic Bottom Water is primarily formed in the Weddell Sea, Prydz Bay, Ross Sea, and Adelie Coast regions (Purkey et al., 2018), which links the properties of Antarctic Bottom Water globally to conditions in these small formation regions. The degree of mixing of the exported Antarctic Bottom Water is unclear from observations (Purkey et al., 2018), but high-resolution modelling shows the export is split by the topography of Drake Passage and Kerguelen Plateau to form distinct Weddell–Prydz-sourced and Ross–Adelie-sourced mixtures in the Atlantic–Indian and Indian–Pacific, respectively (Solodoch et al., 2022). This result suggests that regional changes in Dense Shelf Water formation could produce planetary-scale contrasts in Antarctic Bottom Water properties, and associated changes in the three-dimensional structure of the global overturning circulation.

The upper overturning circulation occurs largely via flow travelling along upward sloping isopycnals in the Southern Ocean interior (J. Toggweiler & Samuels, 1995) meaning that there is no need for a change in the density of the upwelling water mass (the interior flow is largely adiabatic). Diabatic processes occur predominantly at/near the ocean surface, where there is active surface buoyancy and meltwater forcing and wind-driven mixing. However, diabatic transformations are a fundamental necessity to the interior abyssal overturning, as the northward flowing Antarctic Bottom Water must reduce its density and upwell across stable (albeit weak) stratification in the abyss before it can return to the sea-surface (Ganachaud & Wunsch, 2000; Talley, 2013). Diapycnal mixing is the main process that lightens water masses in the abyssal ocean, with geothermal heating a secondary contribution accounting for perhaps 20% (Hofmann & Morales Maqueda, 2009; Emile-Geay & Madec, 2009). Thus, the planetary-scale abyssal overturning is supported by turbulent processes at the Batchelor scale (i.e., on the order of millimetres; Munk, 1966; Ferrari et al., 2016). How and where this buoyancy gain occurs is poorly understood, in part because the interaction between these largest and smallest scales is mediated on intermediate scales, notably by eddies (§ 4.1) and internal gravity waves (§ 5.3).

Diapycnal mixing (§ 4.3.2) of Antarctic Bottom Water is thought to primarily occur where abyssal flows encounter rough bathymetry (Bryden & Nurser, 2003; Fukamachi et al., 2010). Observations near Southern Ocean bottom bathymetry find diapycnal diffusivities that are 10–1000 times greater than upper ocean values (e.g., Heywood et al., 2002; Gara-bato et al., 2004; Polzin, Naveira Garabato, Abrahamsen, et al., 2014). This rapid increase of diapycnal diffusivity with depth causes downwelling in the ocean interior, as water mixes rapidly with denser water beneath it and mixes more slowly with lighter water above it. However, this is compensated by upwelling in the bottom boundary layer where the diapycnal diffusivity goes to zero at the seafloor (Stanley, 2013; de Lavergne et al., 2016; McDougall & Ferrari, 2017; de Lavergne et al., 2017; Cimoli et al., 2019; Holmes & McDougall, 2020). Diapycnal mixing is thought to be sustained by breaking internal gravity waves created from two primary sources: barotropic tides and lee waves resulting from currents interacting with rough topography (§ 5.3.1). Estimates of the amount of Antarctic Bottom Water upwelling driven by tides and lee waves ranging from 7–25 Sv (Nikurashin & Ferrari, 2013; Melet et
Figure 7. Neutral density at the ocean bottom calculated from global hydrographic climatology. The 4 km bathymetric contour is shown in black. The neutral density of Antarctic Bottom Water decreases on its northward journey through diapycnal mixing. Figure reproduced from de Lavergne et al. (2022).

Meanwhile, geothermal heat fluxes are estimated to sustain roughly 2–6 Sv of the abyssal flow (Hofmann & Morales Maqueda, 2009; Emile-Geay & Madec, 2009). These two upwelling effects are offset by a net downwelling associated with the non-linear equation of state of seawater of 6–10 Sv, occurring primarily in the Southern Ocean (Klocker & McDougall, 2010). For the purposes of rough comparison, assuming that the mixing and geothermal upwelling occurs north of 30°S and the non-linear equation of state driven downwelling occurs south of 30°S, gives a mass flux of 9–31 Sv, which is broadly consistent with the maximum northward Antarctic Bottom Water transport of 20—30 Sv near 30°S estimated from observations (Talley, 2013).

Accounting for multiscale processes can alter our fundamental understanding of the dynamics of the abyssal overturning circulation, such as its response to changing the westerly winds over the Southern Ocean. The classic view is that stronger Southern Hemisphere westerly winds, by steepening Southern Ocean isopycnals and altering the abyssal stratification, should weaken the abyssal overturning (Ito & Marshall, 2008; Nikurashin & Vallis, 2011; Shakespeare & Hogg, 2012). However, there is an energetic pathway through which some of the extra wind energy input at the surface leads to enhanced diapycnal diffusion in the abyss, thereby strengthening the abyssal overturning. Specifically, stronger winds steepen isopycnals, driving more baroclinic instability and stronger mesoscale eddies. In the Southern Ocean, these mesoscale eddies are deep-reaching and lead to larger eddy bottom velocities that interact with rough bottom topography to generate lee waves and, ultimately, diapycnal mixing that strengthens the abyssal overturning (D. P. Marshall & Naveira Garibato, 2008; Saenko et al., 2012). When this energetic link is included in idealized models, stronger Southern Ocean westerly winds can actually drive a stronger abyssal overturning (Stanley & Saenko, 2014). Using an estimated climatology of wave energy fluxes, Melet et al. (2014) also found that accounting for lee wave-driven mixing accelerates the abyssal overturning in a realistic global ocean model.
Figure 8. Schematic of the oceanic margin of the southern cryosphere, including key dynamic connections with the Southern Ocean. The ice shelf is the floating extension of the Antarctic Ice Sheet formed from multiple glaciers flowing onto the ocean surface that fuse in suture zones. The ice shelf contains features, such as a melt pond at its surface (that can result in hydrofracture), crevasses and meltwater channels at its base, and rifts that extend throughout the shelf depth and propagate to the shelf front to calve tabular icebergs, from which bergy bits break off. Here, the giant ice shelf partially encloses a cold-water cavity that experiences Mode One circulation, involving bottom inflow of cold water fed by dense shelf water created in a polynya, and outflow of basal meltwater that exits the cavity as a plume of Ice Shelf Water (ISW). At the ice shelf grounding zone, subglacial discharge of ice sheet meltwater flows into the cavity, which creates platelets that attach to the underside of local sea ice. The shelf front is occupied by a polynya (created by katabatic winds) and immobile landfast sea ice (attached to the shelf). Drift sea ice bounds the polynya and landfast sea ice to the north. It is predominantly consolidated sea ice with features, such as pressure ridges, leads and fractures, but with a shear zone at its boundary with the landfast sea ice and a marginal ice zone at its boundary with the open ocean, where floe sizes are relatively small due to the presence of surface waves. Large-scale sea ice drift is dictated by winds, such as those during polar cyclones, as well as ocean currents and gyres.

3 Cryosphere

A major characteristic of the Southern Ocean is that its waters interact with segments of an icy shell created from both freshwater and salt water, respectively, ice shelves and sea ice (Fig. 8). Ice shelves are the thick floating extensions of the land-based Antarctic Ice Sheet onto the coastal ocean surface, which effectively enclose cavities of ocean water. Sea ice is the thin frozen surface of the ocean, which increases the albedo of the ocean surface and, therefore, plays a major role in cooling the Southern Ocean. The sea ice cover is divided into four sub-regions (landfast sea ice, shear zone, consolidated sea ice pack, and marginal ice zone). Ice shelves and sea ice, along with icebergs and polynyas, form an oceanic margin of the southern cryosphere that interacts with the Southern Ocean at a number of scales, ranging from large-scale circulation down, via many processes, to the small diffusive-viscous scales influencing melt and dissolution rates. In this section, the dynamic
interactions between the cryosphere and the Southern Ocean are divided according whether they impact ice shelves and sub-(ice-)shelf cavities (§3.1) or the sea ice cover (§3.2). The dynamic processes included in each sub-section are ordered from the largest to smallest scales.

There are several existing monographs on sea ice, such as Weeks (2010) and Leppäranta (2011), along with collections of reviews, such as D. N. Thomas (2017), including its dynamic interactions with the ocean, although often focused on Arctic sea ice. In addition, there are review articles and collections on specific components of sea ice, including its rheology (Feltham, 2008), its engineering properties (Timco & Weeks, 2010), landfast sea ice (Fraser et al., 2022), and marginal ice zone dynamics (Bennetts, Bitz, et al., 2022b). Wadhams (2000) is a monograph covering both sea ice and icebergs and their role in the climate system. Morales Maqueda et al. (2004) and Smith Jr and Barber (2007) review polynyas. In contrast to sea ice, there is little synthesis information on ice shelves and ice shelf cavities, other than in the context of numerical modelling (M. Dinniman et al., 2016) or basal melt (Burgard et al., 2022).

### 3.1 Ice shelves and sub-ice shelf cavities

Ice shelves (and ice tongues) comprise many merged glacial flows fused together in suture zones (Fig. 8) as the ice continues its oceanward movement. Each ice shelf forms a floating lid on a cavity of ocean water, which is bound on the land side at the “grounding zone” where the ice sheet leaves the land and begins to float. The cavity’s oceanward open boundary is beneath the “shelf front”, i.e., the terminal face of the ice shelf, which is typically a sharp vertical wall formed by calving of icebergs from the ice shelf (Fig. 8). The
Figure 10. Idealised modes of cavity circulation (after Jacobs et al., 1992) and the influence of a polynya. The modes are shown together for convenience but do not necessarily co-exist nor is there a substantial amount of direct observation of these modes. In addition, this elevation view does not capture the horizontal circulatory structure. The shelf front wedge is a buoyant front associated with summer warming that interacts with the Mode Three circulation (Malyarenko et al., 2019). There is also seasonal variability in some of this structure, as polynya formation generally occurs outside of the middle of summer.

Ice–ocean basal interface is the upper boundary of the cavity, where melting and re-freezing takes place.

Ice shelves cover a range of scales. The Ross and Filchner-Ronne Ice Shelf systems are both very large, with areas of the order of 500,000 km$^2$ (Fig. 9). There are intermediate size shelf systems, such as the Amery and Larsen C Ice Shelves (Fig. 9). Smaller shelf systems are very common, such that > 70% of the Antarctica’s grounded ice-sheet boundary sustains an ice shelf of some scale (Bindschadler, Choi, et al., 2011). The ice forming these shelves is typically many hundreds of metres thick (up to 2.5 km in the case of the Amery Ice Shelf). The sub-shelf water cavity can be as much as 1600 m deep (below the Amery). Ice shelf cavities are classified as either “cold” or “warm”, based on the presence (cold) or absence (warm) of water well above the local freezing point (typically Circumpolar Deep Water) in the cavity (Joughin et al., 2012).

3.1.1 Ice shelf cavity exchange with the Southern Ocean

Ice shelf cavities are unique ocean environments. The ocean within a cavity is sheltered from direct wind forcing, but subject to tides and buoyancy induced exchange at the cavity front, as well as the buoyancy fluxes at the ice base due to melting/freezing and heat loss to the ice. Given the importance of mass loss through ocean-driven melting at the base of the ice shelves, it is important to quantify the spatial distribution of melting. The rate and distribution of melting is determined by a complex set of processes that start with the transport of ocean heat into and within sub-shelf cavities.
Cavity circulation, whilst fundamentally three dimensional, generally involves the inflow of dense water at depth (“bottom inflow”; Fig. 8) and outflow of buoyant, meltwater-modified water at the ice base (“basal outflow”). The pathways and preconditioning of inflowing water masses by processes over the continental shelf are unique to each ice shelf (e.g., Nakayama, Timmermann, et al., 2014), and can vary substantially on seasonal, interannual and longer time scales, leading to a wide range of ocean heat inputs (Pritchard et al., 2012; Adusumilli et al., 2020), typically as a bottom inflow. Uncertainties in knowledge of bathymetry and ice shelf base depths create substantial uncertainty in cavity circulation (Greenbaum et al., 2015; Jendersie et al., 2018; Tankersley et al., 2022). Cavity exchange is influenced by water column stratification in response to the circulation (e.g., see §2.2), which compounds the uncertainty (Wåhlin et al., 2020).

Properties of inflowing water masses can experience large seasonal variability (Tinto et al., 2019), which can be reflected in associated basal melt rates, i.e., the rate at which ice is converted to liquid water (e.g., Bindschadler, Vaughan, & Vornberger, 2011; C. L. Stewart et al., 2019). Further, interannual changes associated with climate modes, including the El Niño/Southern Oscillation and the Southern Annular Mode, can also change basal melt rates, primarily through the effects of altered wind fields on ocean heat flows across the continental shelf (Paolo et al., 2018; Adusumilli et al., 2020).

Water mass exchange between the Southern Ocean and ice shelf cavities is typically divided into three modes of circulation (Fig. 10) resulting in the cold or warm cavity descriptor (Jacobs et al., 1992; Silvano et al., 2016). The giant cold cavities of the Filchner-Ronne, Ross and Amery Ice Shelves span hundreds of kilometres across and are typically dominated by Mode One circulation. In this situation, katabatic winds (cold, dense air masses flowing off the polar plateau; L. Thompson et al., 2020; Gutjahr et al., 2022) drive sea ice production in coastal polynyas (§4.2.2) at the ice shelf front. This creates dense shelf water, which floods the cavity and ensures relatively low average melt rates, with some areas of the shelf underside re-freezing (Galton-Fenzi et al., 2012). In addition, this circulation provides protection from warm water inflow (Hattermann et al., 2021; Darelius et al., 2016). Results from smaller shelves, such as the Nansen (Friedrichs et al., 2022) and Sorsdal (Gwyther et al., 2020) Ice Shelves, indicate cold conditions and possibly Mode One circulation can also exist at these scales.

Smaller cavities cover reduced ocean volumes, but, nevertheless, depending on location, can influence substantial terrestrial catchments. For example, while the warm Pine Island and Thwaites Ice Shelf systems are three orders of magnitude smaller in cavity area compared to the giants, they sit at the nexus between a warmed ocean and a significant ice sheet catchment system (Dotto et al., 2022). Similarly, the modest in scale Fimbul Ice Shelf is fed by the East Antarctic Ice Sheet via the Jutulstraumen Glacier (Sinisalo et al., 2013). Because of this, these systems are currently the focus of intense research activity.

Warm water cavities lack the protection of wide, shallow continental shelves, so that (relatively warm) Circumpolar Deep Water has direct access to the underside of the ice shelves. Warm water cavities typically sustain Mode Two circulation (Fig. 10), whereby the inflow of Circumpolar Deep Water leads to high melt rates deep within the cavity. Ice shelves of the Amundsen and Bellingshausen Seas (e.g., Thwaites, Pine Island, Dotson, Crosson and Getz Ice Shelves) are particularly vulnerable and have been observed to have the highest basal melt rates around Antarctica (Rignot et al., 2013; Adusumilli et al., 2020).

Ocean variability in the Amundsen Sea is primarily controlled by winds. Wind anomalies on the shelf break control the import of Circumpolar Deep Water onto the continental shelf (Heywood et al., 2016; P. R. Holland et al., 2019; Dotto et al., 2020), where its flow toward glaciers is controlled by bathymetry (Nakayama et al., 2019). Short-term variability of cavity inflow and associated basal melt rates is controlled by winds in front of ice shelves, and thermocline depth response to this local forcing (Davis et al., 2018). While there is much focus on the current high melt rate of warm cavities, continued warming and freshen-
ing of Dense Shelf Water or enhanced flux of Circumpolar Deep Water onto cold continental shelves may lead to warming of the continental shelf waters and shift the cavities to warm conditions by enhancing Mode Two circulation (Hellmer et al., 2012; Naughten et al., 2021).

Mode Three cavity circulation is associated with the melting that results from an accumulation of warm water along the shelf front. This tends to be more variable than the other modes. In the Amundsen Sea region, Mode Three circulation is associated with Circumpolar Deep Water circulation near the ice shelf front (Davis et al., 2022). Recent observations from the frontal region of the Ross Ice Shelf cavity have shown evidence of high melt rates caused by surface water inflow in the frontal zone directly connected with summer surface ocean warming (C. L. Stewart et al., 2019; Aoki et al., 2022). This buoyant water can potentially pool against the shelf terminal face and form a blocking "wedge" that can influence how waters offshore of the wedge are advected beneath the ice shelf (Malyarenko et al., 2019). While localised, this circulation mode can still have a profound effect on the entire shelf system, depending on where the warming is happening. For example, increased melt rates near Ross Island influence the flow rate of the entire ice shelf (Reese et al., 2018). Meltwater from these ice shelves moves westward in the Antarctic Slope Current (§2.2), and can affect vertical mixing, sea ice production and downstream cavities (Silvano et al., 2018).

3.1.2 Cavities, gyres and eddies

The three modes of cavity circulation (§3.1.1; Fig. 10) need to be augmented with improved understanding of mesoscale variability. Here the literature uses terms like gyre and eddy inconsistently. The terms describe rotating coherent two-dimensional motions with gyres being larger and relatively stationary compared to the smaller eddying motion. These structures have been observed to influence cavity–open ocean exchange, whereby the circulation and associated influence on mixing increases the heat flux into the cavity, thus enhancing basal melting and ultimately resulting in greater freshwater flux into the ocean. This is seen in both warm cavities (A. C. Naveira Garabato et al., 2017; Yoon et al., 2022) and cold cavities (Friedrichs et al., 2022). The Pine Island Glacier ice shelf is a warm cavity example, which shows a system dominated by a gyre that fills the bay in front of the glacier. There is an approximately 10 km wide buffer zone between the shelf front and the gyre, which A. M. Thurnherr et al. (2014) suggests contained ice cavity water.

Friedrichs et al. (2022)’s observations of a small cold cavity (below the Nansen Ice Shelf) also identified the presence and influence of eddies associated with regional topography, including the large Drygalski Ice Tongue, and found that the circulating structure depended on the wind forcing. The authors suggest that the heat being transported into the cavity was influenced by the position of the eddies, both relative to the ice front and to any local basal channels, as it could act as a pump for moving warm water into the cavity.

While these topographically-influenced gyres are relatively large (several tens of kilometres in scale) and stationary, in large cavities and/or away from direct topographic control, eddies are smaller and free to move. Freely moving eddies are typically at the scale of the local Rossby radius of deformation at which rotation effects are comparable to buoyancy effects, which is typically around a few kilometres. At present, numerical modelling is the only way to examine eddy processes within cavities. Mack et al. (2019) quantified the proportion of eddies within a large cavity, along with their scales and persistence. The modelling suggests the eddies persist into the cavity but do not significantly alter the shelf-averaged heat flux.

There are only a few direct observations of the ocean within ice shelf cavities, which can be used to assess models. Some borehole data are providing information at specific locations (Stevens et al., 2020). More recently, robotic technology is providing a view of the vertical structure and its spatial variations (Fig 11; Gwyther et al., 2020; Graham et al., 2022; Davis et al., 2022). Data of this type provides the first evidence of the mixing and diffusivity in these under-observed environments. This is particularly important because
of the often long circulation timescales (several years in some cases) within cavities, and limited set of drivers. These conditions mean that even modest differences in ocean cavity turbulence results in sustained diffusivities and associated heat flux. This is in contrast to boundary-driven mixing in the Southern Ocean with many coincident driving processes (§4.3.3).

3.1.3 Tidal influence on cavities

Due to the absence of direct weather forcing within a cavity, ocean tides (both internal and surface; §5.2) are the primary forcing at periods within the 0.5–10 day range. The elastic response of an ice shelf to any large-scale perturbation means that, other than close by the shore at the grounding zone, the ice responds hydrostatically and rapidly. Thus, determination of tidal excursions and currents can be achieved in the same way as elsewhere in the oceans, by combining water column height observations, knowledge of the bathymetry and numerical tools to extrapolate to any location in space and time (Padman et al., 2018). There are subtleties to tidal mechanics at such high latitudes, as the influence of tides on ice shelf melting is related to the latitude of an ice shelf relative the semi-diurnal critical latitude, where the tidal frequency equals the inertial frequency (§5.3.3; Robertson, 2013). In general, tides can be important drivers of meltwater production for ice shelves on cold water cavities (e.g., Makinson et al., 2011; Arzeno et al., 2014; Mueller et al., 2018; Hausmann et al., 2020), but are less important for ice shelves on warm water cavities (e.g., Robertson, 2013; Jourdain et al., 2020).

Accepted melt rate parameterisations involve the local under-ice velocity (D. M. Holland & Jenkins, 1999; Rosevear, Galton-Fenzi, & Stevens, 2022). However, including tides in
3.1.4 Meltwater plumes and marine ice

In the far reaches of an ice shelf cavity, once the inflowing oceanic water mass comes in contact with the ice shelf, production of meltwater results in a buoyant plume at the ice base underside (Fig. 8). The meltwater plume typically ascends as it travels oceanwards, steered by the ice base topography and coastlines, and drives cavity-scale convective circulation. The evolution of the meltwater plume is governed by friction, planetary rotation and the entrainment of underlying water masses (Jenkins, 1991). Since the in situ melting point is reduced by approximately $0.75^\circ$C per kilometer of depth, cold water, such as Dense Shelf Water that is typically at the surface freezing temperature, drives rapid melting at depth. For cold cavities, rapid melting at deep grounding zones can lead to potentially “supercooled” plumes that rise along the ice base to a point where in situ freezing occurs — the so-called “ice pump”. This occurs when basal melting at the grounding zone results in a meltwater plume that then can re-freeze at shallower depths (Lewis & Perkin, 1986; Schodlok et al., 2016), sometimes through the formation of platelet ice crystals (Hoppmann et al., 2020). At that point, ice forms and rises to accrete to the ice shelf base as “marine ice” (Stevens et al., 2020). The spatial patterns of melting and refreezing can be seen in satellite altimetry data (Adusumilli et al., 2020). Under warm cavities, not all in-flowing ocean heat is consumed. Instead, the meltwater plume brings ocean heat to the surface and forms near-ice-front sensible heat polynyas (e.g., Mankoff et al., 2012).

Localized plumes of Ice Shelf Water can exit the cavity and be found across the continental shelf (e.g., A. H. Orsi & Wiederwohl, 2009), and can be a substantial contributor to the formation of Antarctic Bottom Water (§2.5) with far-reaching implications for issues such as global sea level rise, the fate of Antarctic sea ice, ongoing changes in bottom water formation and global deep-ocean ventilation, and the evolution of Southern Ocean ecosystems and carbon cycling (A. C. Naveira Garabato et al., 2017; Lago & England, 2019; Hazel & Stewart, 2020). For example, they provide a pathway for iron (a nutrient) from the sediment to the ocean surface to seed phytoplankton blooms (Herraiz-Borreguero et al., 2016). For cold cavities, the ice pump extends beyond the cavity itself and meltwater plumes exiting the cavity along the ice shelf base, or at depth Fig. 8) are often supercooled and directly contribute to sea ice growth (N. J. Robinson et al., 2017; Hoppmann et al., 2020).

The characteristics of the meltwater plumes are influenced by ice base topography, with basal channels being sites of enhanced basal melting (e.g., W. Wei et al., 2020), and the presence of other forcing (primarily tides; §5.2) of turbulent mixing at the ice-ocean interface. Plume circulation and melt rates are expected to be altered by the presence of tides, but the direction and magnitude of the change depends on the balance between tide-enhanced drag, entrainment and melting (Anselin et al., 2023).

3.1.5 Cavity grounding zone and subglacial discharge

The “grounding zone” of an ice shelf cavity is the region where the terrestrially-borne ice lifts off from the bedrock and sediment, and encounters the ocean (Fig. 12b). While some-
Figure 12. Small-scale views of an ice shelf and sub-shelf water cavity showing some of the under-observed but critical processes likely to be present. (a) Grounding zone region including subglacial discharge of meltwater from beneath the ice sheet, basal crevasses, stratification/baroclinic waves (dashed lines), in/outflow. (b) The basal boundary layer (bbl), temperature and salt stratification and roughness variations. The basal boundary layer shows the 1–10 m thick region close to the ice-shelf base, where the fluid velocity and turbulence is affected by the presence of buoyant meltwater. Temperature and salinity increase rapidly through the boundary layer from diffusion-controlled melting conditions at the ice–ocean interface through to the boundary layer itself and then to ocean-cavity conditions at the edge of the boundary (the circulation of which is not well known).

Times referred to as a “grounding line”, key ocean interactions happen over a wide region. Beneath hundreds-to-thousands of metres of ice, this is one of the most challenging ocean environments to observe directly, so most inferences come from surface-based measurements (e.g., Jenkins et al., 2006) or remote sensing (Rignot et al., 2011).

Surface detection of the grounding line provides some guidance as to the nature of the region where the ice lifts from the bed and, in particular, the importance of tidal effects (Wild et al., 2018; Fricker, Allison, et al., 2002). The lift-off is potentially modulated by tidal and pressure effects, as observed in the WISSARD borehole (Begeman et al., 2020) at the southernmost extent of the Ross Ice Shelf. The vertical excursion potentially influences ice shelf flow, so that the tidal modulation enables the ice to be advanced oceanward (Rosier & Gudmundsson, 2020; Drews et al., 2021).

Subglacial discharge is the flow of meltwater from the ice sheet basal bedrock interfacial zone that finds its way into the cavity coastal zone (Fig. 12a). The meltwater is formed by pressure and geothermal warming (Fricker et al., 2016). While difficult to access in Antarctica, subglacial discharges of meltwater have been extensively studied in the context of the Greenland Ice Sheet, where they are often linked to elevated melt rates (I. J. Hewitt, 2020). Satellite observations provide evidence for a large number of active subglacial lakes across Antarctica, which experience sporadic but rapid drainage events (Fricker et al., 2007; Siegfried & Fricker, 2018). As in Greenland, it is likely that these drainage events alter the water properties at the edge of the ice sheet and have a significant effect on ice–ocean processes (Miles et al., 2018; Jouvet et al., 2018). However, the nature, frequency, and location of these subglacial drainage events remain unclear, largely due to challenges in making oceanographic observations deep within the ice shelf cavity.
Figure 13. Comparisons of observed to predicted melt rates for a variety of ice shelf situations. The symbol colours mark the ratio of observed ($m_{\text{OBS}}$) to predicted melt rate based on the parameterisation by Jenkins, Dutrieux, et al. (2010) ($m_{\text{J10}}$), as a function of thermal driving ($T_*$) and free stream velocity ($U$). The underlying contours showing the predicted $m_{\text{J10}}$ melt rates in m yr$^{-1}$. The map shows the location of the observations. Figure reproduced from Rosevear, Galton-Fenzi, and Stevens (2022).

The injection of freshwater, either from sub-glacial discharge or from basal melt, causes the water column near grounding lines to exhibit aspects of an estuary with landward-flowing deeper water exchanging with this freshwater flux (Horgan et al., 2013). The resulting stratification is influenced by tidal mixing processes through both mixing and baroclinic waves (Fig. 12a), with a key question being at what point does the tidal mixing become sufficient to homogenize the water column (P. R. Holland, 2008). The few observations available suggest stratification can persist in even quite thin water columns (e.g., 10–30 m; Begeman et al., 2018; Lawrence et al., 2023; Davis et al., 2023). This suggests that inflowing warm water can directly access the basal boundary layer right at the formation of the ice shelf meltwater plume.

3.1.6 Cavity basal boundary layers

The basal boundary layer is the oceanic boundary layer just beneath the base of an ice shelf (Fig. 12b). The assumption is it is a turbulent boundary layer but given likely slow flow speeds, smooth boundary and strong stratification due to the meltwater flux, this may not be true everywhere. The flow within the boundary layer is responsible for setting the ice shelf basal melt rate, which affects both ice sheet stability and ocean circulation, by regulating heat and salt transport from the ocean to the ice via turbulent processes on small scales (of order centimetres to metres). It also drives the basal meltwater outflow from the cavity.

The archetypal model of the ice shelf basal boundary layer (Fig. 12b) is of a boundary layer formed by velocity shear due to friction between the ice and ocean currents. These currents may be buoyant meltwater plumes, tidal currents, eddies, or other mean circulation within the cavity (Stanton et al., 2013; Padman et al., 2018), all of which are poorly known in cavities (Fig. 12b). In this “shear-driven” regime, the basal melt rate depends
on the friction velocity (a turbulent velocity scale related to the current speed) and ocean temperature (Davis & Nicholls, 2019; Vreugdenhil & Taylor, 2019; Rosevear, Gayen, & Galton-Fenzi, 2022). This model forms the basis of common ice–ocean parameterizations (e.g., D. M. Holland & Jenkins, 1999; Jenkins, Nicholls, & Corr, 2010). However, comparisons between observed and predicted melt rates of ice shelves, as well as idealised models, have brought into question the appropriateness of this approach when current velocities are low (Malyarenko et al., 2020; Rosevear, Galton-Fenzi, & Stevens, 2022) or the near-ice stratification is strong (Vreugdenhil & Taylor, 2019). The departure of observed ice-shelf melt rates from those predicted by a standard ice–ocean parameterisation is shown in Fig. 13 for a range of ice shelves, spanning cold and warm cavities, with the departure in some cases reaching nearly an order of magnitude. This also gives a sense of how limited our sampling of this key comparison is. Essentially the same small-scale melting processes are parameterised differently when applied to the melting of icebergs (§3.1.7; Hester et al., 2021).

Meltwater is less dense than ambient seawater, primarily due to salinity differences, and will drive convection in the form of a buoyant plume if the ice-shelf base is sloped (Figs. 8, 12). This gives rise to a convective melting regime (seen in laboratory experiments and simulations), in which melting is driven by gravitational instability (Kerr & McConnochie, 2015; McConnochie & Kerr, 2017b; Gayen et al., 2016). Convective melting is independent of the current speed and, instead, depends upon temperature and the slope of the ice base, with the most rapid melting occurring for vertical ice (McConnochie & Kerr, 2018; Mondal et al., 2019). Antarctic ice shelves typically have low slope angles, which inhibits the gravitational instability. Thus, convective melting may be more relevant to near-vertical ice such as icebergs and shelf fronts. A transition from convective- to shear-driven melting is expected as the buoyant plume gains momentum (Malyarenko et al., 2020; McConnochie & Kerr, 2017a). However, this transition is poorly constrained and may vary over quite small scales (Schmidt et al., 2023a). A general description of this important boundary condition has yet to be derived.

The role played by buoyant meltwater depends on whether the ice base is sloped or horizontal, and what other forcing is present. For a shear-dominated boundary layer beneath horizontal ice, meltwater is expected to stratify the boundary layer and suppress turbulence. Recent numerical simulations have shown that buoyancy inhibits melting by decreasing the efficiency of heat and salt transport to the ice boundary (Vreugdenhil & Taylor, 2019) and insulating the ice from warmer water below (Rosevear, Gayen, & Galton-Fenzi, 2022). When shear is weak, the heat and salt fluxes associated with basal melting provide an opportunity for diffusive convection to occur (a type of double-diffusive convection; L. Middleton et al., 2021; Rosevear et al., 2021). Diffusive convection is the process by which a stably stratified fluid becomes unstable due to the large difference in the molecular diffusivities of heat and salt, forming well-mixed layers separated by thin interfaces called “thermohaline staircases” (Radko, 2013). Observations from beneath the George VI Ice Shelf show a persistent staircase (S. Kimura et al., 2015), and weak dissipation, which is uncorrelated to current speed (L. Middleton et al., 2022), suggesting that diffusive-convection is the primary driver of turbulence. There is also evidence of diffusive-convection-susceptible conditions beneath the Ross Ice Shelf (Begeman et al., 2018).

Small scale observations of ice-shelf basal surfaces have shown a variety of features including channels, crevasses and step-like terraces (Fig. 12a; Dutrieux et al., 2014; Stanton et al., 2013; Lawrence et al., 2023; Schmidt et al., 2023a). Feedbacks between ice topography and basal melting are expected due to differential melting of sidewalls versus horizontal surfaces (McConnochie & Kerr, 2018; Mondal et al., 2019) and the possibility of large features such as channels (Alley et al., 2016; Gourmelen et al., 2017; Chartrand & Howat, 2020) and crevasses (Lawrence et al., 2023) guiding flow, especially meltwater plumes and boundary currents, which are relatively thin (~10 m). Smaller-scale basal texture or “roughness” (Fig. 12b) is expected to enhance boundary-layer turbulence, leading to higher melt rates (Gwyther et al., 2015), and sapping momentum from buoyant plumes through increased
drag (e.g., Smendsrud & Jenkins, 2004). There are very few direct measurements of turbulence or drag beneath ice shelves (Stanton et al., 2013; Davis & Nicholls, 2019; Venables et al., 2014; L. Middleton et al., 2022). This is in part because the borehole can affect the boundary layer making undisturbed measurement challenging. Autonomous vehicles are providing a platform that circumvents this challenge (Davis et al., 2022). Beneath the warm Larsen C Ice Shelf, a relatively low drag coefficient was observed (Davis & Nicholls, 2019). However, sea-ice analogs for marine ice zones (refreezing regions formed by the accretion of frazil ice) suggest that drag coefficients up to two orders of magnitude higher are possible (N. J. Robinson et al., 2017).

3.1.7 Iceberg calving

About half of the mass loss from the Antarctic Ice Sheet is via icebergs calving from the shelf fronts (Rignot et al., 2013; Depoorter et al., 2013; Greene et al., 2022). Mass loss from the three largest ice shelves (Ross, Filchner-Ronne and Amery), which cover cold-water cavities (§3.1.2), is dominated by calving (Lazzara et al., 1999; Fricker, Young, et al., 2002; Walker et al., 2021). Icebergs range from massive tabular icebergs, with areas sometimes exceeding 1000 km$^2$, to small “bergy bits” formed by erosional processes at ice fronts and edges of larger icebergs (Fig. 8). Most mass loss due to calving is via large tabular icebergs, which tend to calve from any particular source region on decadal time scales. These large icebergs can persist for decades. This contrasts with the melting of smaller icebergs, which happens more rapidly and forms a distributed source of freshwater along the coastal margins of the continent (Merino et al., 2016).

Ice shelf calving events are a consequence of the propagation of rifts (crevasses that penetrate the full shelf thickness) to the shelf front, such that they isolate ice blocks from the main shelf (an anticipated calving site is represented in Fig. 8). Spatial variations in ice shelf velocity are the ‘first-order control’ on calving, as they cause strain rates that determine the location and depth of crevasses and, subsequently, propagate the crevasses and resulting rifts (Benn et al., 2007). These phenomena occur at the scale of the ice shelf flow structure (Meier, 1997). However, smaller scale processes are also present, such as “hydrofracturing”, where the water in surface melt ponds flows into and expands surface crevasses can have significant influence on an ice shelf’s resilience (Fig. 8; Lai et al., 2020).

Once crevasses or rifts have formed in ice shelves, force imbalances due to the surrounding water also drive crevasse and rift propagation (Benn et al., 2007). Hence, dynamic couplings between ice shelves and the Southern Ocean exert important ‘second-order controls’ on iceberg calving (Y. Liu et al., 2015). There is evidence that this only occurs once the ice shelf has thinned sufficiently or for a rift system that is close to detachment (Bassis et al., 2008). Moreover, if present, fast ice or mélangé (a consolidated agglomeration of icebergs and fast ice) exerts a backstress on ice shelves (Massom et al., 2010; Greene et al., 2018), which can delay or prevent iceberg calving (Stevens et al., 2013; Massom et al., 2015, 2018; Arthur et al., 2021; Gomez-Fell et al., 2022).

In addition to these slowly varying drivers of iceberg calving, there are wave-driven mechanisms of relevance also. Ice shelf flexure has been detected in response to surface waves ranging from long-period swell (§5.1.1) to infragravity waves and tsunamis. Flexure due to swell is greatest in the outer shelf margins (Chen et al., 2018; Bennetts, Liang, & Pitt, 2022) and during summer when the sea ice barrier is at its weakest or absent (Massom et al., 2018; Chen et al., 2019). Swell-induced shelf stresses peak at crevasses (Bennetts, Liang, & Pitt, 2022), and they have been associated with crevasse and rift propagation (Banwell et al., 2017; Lipovsky, 2018), iceberg calving (MacAyeal et al., 2006; Cathles IV et al., 2009) and triggering catastrophic ice shelf disintegration events (Massom et al., 2018). Infragravity waves and tsunamis are not attenuated by sea ice and cause year-round ice shelf flexure (Bromirski et al., 2017), which can extend long distances onto shelves (Bromirski et al., 2017).
Their impacts have been implicated in calving and disintegration events (Bromirski et al., 2010; Brunt et al., 2011).

Surface waves also initiate small-scale calving through a combination of warm surface water and forced convection. The combination of conditions causes a relatively high rate of melting at the shelf front waterline and a so-called “wavecut”. The wavecut isolates the overhanging ice, which becomes unstable and collapses (Orheim, 1987; T. Hughes, 2002), leaving behind a protruding “ice bench” (or “ice foot”) at the shelf front. The bench exerts a buoyant vertical force, deforming the shelf front into a so-called “rampart moat” structure. The associated internal ice stresses can propagate basal crevasses and, hence, calve relatively small, but full-thickness icebergs along the crevasse.

3.2 Sea ice cover

Sea ice cover is divided into four seasonally changing zones: (i) the largely immobile landfast sea ice (or fast ice), which is attached to many stretches of the coastline, including ice shelf fronts; (ii) the shear zone that sits between the coastline/fast ice and (iii) the consolidated ice pack; and (iv) the highly dynamic outer tens to hundreds of kilometres of the ice cover, known as the marginal ice zone, which is characterised by the presence of surface waves (Fig. 8). Sea ice forms a nearly-continuous ring (interrupted by the Antarctic Peninsula) around Antarctica, which is typically quantified in terms of areal concentration of ice cover on the ocean surface (Fig. 9), as well as the ice thickness and floe size distribution. The ring grows to 1000 km wide during winter, at which point it roughly doubles the area of the southern polar cap (Fig. 9), and shrinks to almost zero during summer. Therefore, the majority of Antarctic sea ice is less than one year old, so that it grows only up to tens of centimetres in thickness, unless it is mechanically deformed, for example into pressure ridges (Fig. 8). Sea ice initially forms as new ice (frazil, grease, slush or shuga), develops into nilas (a thin elastic crust) or pancake ice (disks of sea ice, tens of centimetres to a few metres in diameter, which typically form in wavy conditions), and finally consolidates into chunks known as floes, which range from tens of metres to tens of kilometres in diameter. Floe sizes are typically tens to hundreds of metres in diameter in the marginal ice zone, due to wave-induced breakup of larger floes, and separated by open water, frazil, brash (small bits of ice wreckage) or other interstitial ice. In winter, pancakes (which are often considered as small floes) separated by thin interstitial ice occupy large expanses of the marginal ice zone. There is a sharp increase in mean floe sizes between the marginal ice zone and the consolidated ice pack, accompanied by the absence of persistent wave activity. Floe sizes in the consolidated ice pack can be large enough (kilometres or more in diameter) to consider the ice cover as a quasi continuum, punctuated by ridges, cracks and leads, as well as polynyas and embedded icebergs (Fig. 8).

Polynyas are an additional phenomenon that form around the Antarctic margin. They are non-linear openings in the ice cover (i.e., not leads or fractures) where sea ice would be expected for thermodynamic reasons alone. They are created by local melting of sea ice by warm water upwelling or katabatic winds driving sea ice away from a location, or a mixture of these two processes. Polynyas range from small, ephemeral polynyas, through to the very large Ross Sea and Cape Darnley polynyas, as well as the open ocean Maud Rise polynya. A contemporary review of polynyas is given in §§ 4.2.2–4.2.3, rather than in § 3.2, due to their important influence on turbulent convection.

3.2.1 Sea ice drift

Sea ice away from the coast, islands or icebergs (where it is usually found as fast ice) is able to drift under forcing from the atmosphere and ocean, and is known as “drift ice” or “pack ice”. Variability in drift redistributes the pack, and influences ice extent, concentration, and thickness and floe size distributions. It also interacts with important components
Mean drift speeds of Antarctic sea ice are \( \approx 0.1 \text{ m s}^{-1} \) (Heil et al., 2006). Generally slower speeds are found in the shear zone, where sea ice predominantly drifts westwards due to easterly coastal winds (Emery et al., 1997; Heil & Allison, 1999; N. Kimura, 2004). The Antarctic Divide separates the near-continuous westward drift around the coastline from eastward drift at lower latitudes that follows the Antarctic Circumpolar Current (§ 2.1). The eastward drift is typically faster as the ice is close to free drift (P. R. Holland & Kwok, 2012). The fastest drift speeds are found in the unconsolidated outer margins of the ice cover (Heil & Allison, 1999; Doble & Wadhams, 2006), where the fastest reported speed is \( \approx 0.75 \text{ m s}^{-1} \), for pancake ice in the marginal ice zone during a winter cyclone (Alberello et al., 2020). In contrast, floe–floe interactions in the shear zone, for example, grinding, rafting and locking, create internal stresses that slow the drift. The westward coastal shear flow is disrupted by intervals of northward ice drift across the Antarctic Divide, where the ice in the shear zone becomes entrained into the inner ice pack (Heil & Allison, 1999). The most prominent regions of northward ice transport are the near cyclonic ice drift patterns that coincide with the Weddell and Ross Gyres (§ 2.3) and the Prydz Bay Gyre at the termination of the Amery Ice Shelf around 80\(^\circ\)E (Emery et al., 1997; N. Kimura, 2004). The proximity to the coast means these drift patterns are potentially influenced by katabatic winds (Emery et al., 1997).

From the perspective of climate, there is a focus on circumpolar Antarctic sea ice coverage metrics. However, being essentially a two phase interfacial phenomenon, a number of small scale processes are directly relevant to the global scale. On time scales of hours or less, atmospheric stress due to winds is generally the dominant driver of sea ice drift (P. R. Holland & Kwok, 2012), with the motion opposed by oceanic stress. Both atmospheric and oceanic stresses are usually modelled using quadratic drag laws, with constant drag coefficients and turning angles (also known as Ekman angles) that represent the difference in direction between the geostrophic flow and the stress on the ice surface due to the Coriolis force (counter-clockwise in the Southern Hemisphere; Weeks, 2010). The drag coefficients can be decomposed into viscous “skin” drag and “form” drag, where the latter depends on the ice roughness, created by an accumulation of relatively small-scale features, particularly floe edges in the marginal ice zone and pressure ridges in the inner ice pack (Tsamados et al., 2014).

Atmospheric and oceanic drag manifest from similar underlying physics, with surface layers above and below the ice, where the stress is approximately constant, followed by Ekman layers, in which the stress decays to zero whilst the velocity rotates due to the Coriolis force (Leppävirta, 2011). The surface layers are typically turbulent, so the drag coefficients are largely independent of relative velocities. However, as typical ice drift speeds are much slower than wind speeds but comparable to ocean current speeds, the wind acts as an external force, whereas ice and ocean dynamics are coupled. The coupled ice and ocean dynamics are dependent on the relative ice velocity, ice basal roughness and the ocean stratification, all under the influence of the Coriolis force (McPhee, 2008). The Coriolis force influences ice drift through the sea surface tilt, which has been attributed as the source of ice drift rotation close to the inertial frequency in water too deep to be caused by tidal currents (Alberello et al., 2020).

### 3.2.2 Surface wave–floe interactions in the marginal ice zone

The marginal ice zone is the region in which sea ice cover is regularly influenced by surface waves that penetrate beyond the ice edge (i.e., aren’t dissipated or scattered away by the transition) and propagate for tens to hundreds of kilometres into the ice-covered ocean before they are attenuated by interactions with the ice cover (§ 5.1.4). Waves affect ice cover in the marginal ice zone (Fig. 14) by (i) breaking up larger floes (see below), (ii) herding...
Figure 14. Schematic of surface wave–ice floe interaction processes in the marginal ice zone, including (from left to right): wave-induced breakup of a large floe; subsequent northward drift of small broken floe due to off-ice winds, and enhanced melt in summer; wave overwash of a floe; herding and rafting of small floes; floe–floe collisions; production of frazil in the open water created between floes during winter.

the floes into bands (Wadhams, 1983; Shen & Ackley, 1991), (iii) promoting growth of new ice (e.g., frazil) during freezing conditions, (iv) forcing ice drift through momentum transfer (radiation stress; T. D. Williams et al., 2017; P. Sutherland & Dumont, 2018; Dumont, 2022), (v) causing floes to collide and raft (S. Martin & Becker, 1987, 1988; Dai et al., 2004; Rottier, 1992; Bennett & Williams, 2015; Yiew et al., 2017; Herman et al., 2019), which gives pancake ice its distinctive raised edges and can erode the floe edges and produce ice rubble, (vi) overwashing the floes (Skene et al., 2015; Nelli et al., 2020; Pitt et al., 2022), which influences thermodynamic ice properties by removing snow cover and creating saline ponds on the floe surfaces (Ackley & Sullivan, 1994; Massom et al., 1997, 2001), and (vii) generating turbulence in the water below floes that increases basal melt. (Wadhams et al., 1979; M. Smith & Thomson, 2019) Overall, the waves create a fragmented, granular ice cover in the marginal ice zone, containing a mixture of floes (smaller than in the inner ice pack) and unconsolidated sea ice (grease, etc.). Ice in the marginal ice zone is highly mobile and responds rapidly to forcing by strong winds over the Southern Ocean (Vichi et al., 2019; Alberello et al., 2020).

Breakup of large floes is usually considered to be the primary effect of waves on sea ice. Ice floes larger than prevailing wavelengths experience a hydroelastic response to wave motion (Montiel, Bonnefoy et al., 2013; Montiel, Bennett, et al., 2013; Meylan et al., 2015), creating so-called “flexural-gravity waves” (Bennetts et al., 2007; Vaughan et al., 2009). Sea ice is a brittle material (Timco & Weeks, 2010), which fractures when the flexural stresses/strains exceed the material strength (Montiel & Mokus, 2022). The generally held view of the wave-induced breakup process (Squire et al., 1995) is of a large wave event breaking up a quasi-continuous sea ice cover (e.g., a very large floe) into smaller floes that then form or join the marginal ice zone, in which floes larger than the prevailing wavelengths are broken up further, thus forming a marginal ice zone where mean floe sizes increase away from the sea ice edge as wavelengths increase (Squire & Moore, 1980). The standard theoretical description of the wave-induced breakup process is of regular (unidirectional
and monochromatic) flexural-gravity waves in a homogeneous floating elastic plate causing stresses/strains that exceed a critical threshold (Kohout & Meylan, 2008; Vaughan & Squire, 2011; Mokus & Montiel, 2021; Montiel & Mokus, 2022). This framework has been extended to irregular (T. D. Williams et al., 2013a, 2013b) and nonlinear (Herman, 2017) waves, and three-dimensional effects (Montiel & Squire, 2017). Contemporary models often use an empirical strain threshold from the small set of historical observations reported by Goodman et al. (1980) and Squire and Martin (1980) (Vaughan & Squire, 2011; Kohout & Meylan, 2008; Horvat & Tziperman, 2015; Mokus & Montiel, 2021) or empirical models based on thousands of flexural strength tests conducted on sea ice sturdy enough for experimental testing (i.e., not in the marginal ice zone; T. D. Williams et al., 2013a, 2013b; Montiel & Squire, 2017). The critical thresholds are similar, albeit those derived from flexural strength tests are slightly larger. Modern field measurements of wave-induced breakup are often conducted on fast ice (Voermans et al., 2019), where it is possible to instrument the ice cover before breakup, or in a “natural laboratory” (in a bay of the Gulf of St Lawrence using ship generated waves; Dumas-Lefebvre & Dumont, 2023). Laboratory experiments in ice tanks are also being used to generate more detailed measurements of the waves and ice during the breakup process (Dolatshah et al., 2018; Herman et al., 2018; Passerotti et al., 2022), although the scaling issues associated with model ice (Schwarz, 1977; von Bock und Polach et al., 2015, 2019) complicate interpretation of the results for breakup in the marginal ice zone. Measuring breakup in the marginal ice zone remains challenging, despite it being identified as a priority three decades ago (Squire et al., 1995).

A positive (summer) feedback between wave-induced breakup and ice weakening (or loss) has been proposed (Bennetts et al., 2010; Montiel & Squire, 2017; Horvat, 2022), in which an initial weakening of the ice cover allows waves to travel farther into the sea ice-covered ocean, so that a wave event can break the ice cover at a deeper location than prior to the initial weakening. The breakup leaves the floes more susceptible to lateral melting during the summer (Steele, 1992), which further weakens the ice cover and allows waves to travel even deeper into the ice-covered ocean, and so on. It has been suggested that the positive feedback has already been triggered in the Arctic due to initial weakening by warming temperatures (Squire, 2011), although there is no evidence of this, as yet. However, the feedback is implicit in the comparison between trends in Antarctic ice edge latitude and local significant wave heights (Kohout et al., 2014). A negative (winter) feedback has also been proposed, in which wave-induced breakup creates openings in the ice cover (leads; Fig. 8) that freeze over to strengthen the ice cover and insulate the location against future wave events (Squire, 2011; Horvat, 2022).

### 3.2.3 Brine inclusions to convective channels

Sea ice is a complex blend of solid H$_2$O crystals, liquid brine inclusions, air bubbles and precipitated salts (Fig. 15), whose volume fractions, distributions and connectedness depend strongly on temperature, salinity and depth (Perovich & Gow, 1996; Light et al., 2003; Golden et al., 2007; Golden, 2009; D. N. Thomas, 2017; Kraitzman et al., 2022). The air bubbles are uniformly and isotropically distributed throughout the ice (Golden, 1998), and they are less elongated and not vertically oriented like the brine inclusions. Air bubbles are about five times larger than brine inclusions, but the volume fraction of air bubbles is only about one-third of the volume fraction of brine inclusions (Perovich & Gow, 1996). Moreover, sea ice usually consists of a layer of “granular textured ice”, above a layer of “columnar textured ice”, separated by a transition layer. The columnar layer is characterised by large, well-ordered ice crystals with vertical orientation, while granular textured ice has random crystal orientation (Eicken, 2003; Lund-Hansen et al., 2020; Oggier & Eicken, 2022). Due to dynamic growth conditions in the Southern Ocean, more than 60% of the total thickness of Antarctic sea ice is primarily composed of frazil ice, and, in the upper layer of the ocean, frazil ice tends to form floes that contain a significant amount of ice with a granular texture. In turn, this leads to the dominant layer of Antarctic sea ice being characterised by a granular texture (Lange et al., 1989; D. N. Thomas, 2017).
Sea ice micro- and fine-structure determines its large-scale (bulk) physical (Trodahl et al., 2001; Pringle et al., 2007) and optical properties (Perovich, 1996; Light, 2010), and both seasonal and spatial variations in its microstructure influence the flaw structure of the sea ice (i.e., imperfections that develop as a result of microphysical processes that occur during ice growth) and its fracture and crack propagation (i.e., larger-scale features that result from subsequent thermal or dynamic effects D. M. Cole, 2001). Antarctic sea ice, with its unique small-scale structure, which is characterised by a relatively high volume of first-year ice and growth under extremely low temperatures, exhibits higher values of thermal conductivity than Arctic sea ice. Moreover, the salinity concentration of the brine inclusions and the precipitation of solid salt crystals govern the sea ice refractive index (Nelson & Thompson, 1954; Richardson, 1976; Marion et al., 1999; Maykut & Light, 1995) and the distribution of liquid brine within the ice matrix and the varying width of solid ice layers control the value of the sea ice albedo. As Antarctic sea ice is mainly young, first-year ice, it is full of liquid brine inclusions and allows large quantities of light to be transmitted to the ocean (Light, 2010).

On the micrometre–centimetre length scale, liquid water molecules solvate the salt ions, and their removal by the freezing process unfavourably decreases the entropy of the ions. The resulting ejection of salt from the ice matrix engenders a chemotactic flow for the salt density that leads to the development of spatially extended regions of high salt concentration, i.e., the brine inclusions (Kraitzman et al., 2022) (Fig. 15 zoom). In the Southern Ocean, the volume fraction of brine inclusions changes seasonally between 2–40% (Jacques et al., 2021). From mid-November to mid-January, the ice melting season takes place in the southern region. Despite the atmospheric temperature not being sufficiently warm to melt the ice, the ocean plays a dominant role in controlling the melting process. During this season, the Southern Ocean experiences considerable heat fluxes to the atmosphere, with values as high as 50 W m$^{-2}$ (Gow & Tucker III, 1990). During the melt season, the brine inclusions expand.

Figure 15. The main panel (left-hand side) is a schematic of sea ice at finescale. Above the dashed line, air bubbles (white circles) and brine inclusions (elongated blue shapes) are trapped within the impermeable, solid ice (sky blue). Below the line the ice is permeable, allowing brine drainage and fresh water inflow, i.e., a mushy layer. A zoom in on the microscale for a brine inclusion is given (right). The liquid brine region is surrounded by ice, and the arrows point in the direction of the salt flux during the freezing process.
and merge to form up to metre-long brine channels, which allow fluid, nutrients and salt to
exchange between the ocean and the ice (Golden et al., 1998; Golden, 2001). For Antarctic
sea ice, brine channels are vertically oriented with diameters $\approx 200 \mu \text{m}$ (Weissenberger et al.,
1992), and the brine fluid flow in the channels is a critical factor in the facilitation of thermal
fluxes, which leads to an enhancement in the thermal conductivity (Lytle & Ackley, 1996;
Trodahl et al., 2001). Moreover, the brine drainage leads to the formation of air bubbles,
which result in greater albedo (Perovich, 1996).

On the centimetre–metre length scale, sea ice is commonly described as a “mushy
layer” (solid ice crystals mixed with interstitial liquid brine), bounded from above by an
impermeable layer and from below by a fully liquid layer (Fig. 15; Feltham et al., 2006).
The dense, salty interstitial fluid is trapped and stagnant within the ice matrix and is
assumed to be in thermodynamic equilibrium, which prevents the solid ice from melting.
As the sea ice grows, the interstitial liquid in the mushy layer undergoes convection due
to differences in temperature and density, leading to the release of salt into the ocean.
This brine drainage phenomenon is accompanied by inflow of less saline seawater from the
surrounding mushy layer (Worster et al., 2000; Worster & Jones, 2015; A. Wells et al., 2011;
A. J. Wells et al., 2019). With a local convective flow partially occupying the mushy layer,
brine drainage occurs in only part of the sea ice. However, as the temperature increases
and the sea ice becomes more porous, the convective flow eventually spreads throughout
the entire sea ice depth, utilising the network of brine channels. Oceanic currents exert
pressure on the interface layer between the sea ice and the ocean, affecting the convective
brine flow (Feltham et al., 2002). Many contemporary models are revising their sea ice
representation to include its complex, evolving microstructure (A. K. Turner et al., 2013;
Hunke et al., 2015). Moving from a static, constant salinity profile representation of sea
ice to a time-evolving and thermodynamic model has modified the state of the sea ice, the
ocean, and the atmosphere, as well as their fluxes (DuVivier et al., 2021). Incorporating
mushy layer theory in climate models has a significant impact on simulation results, such
as an increase in both the snow-ice and the coastal ice production, and a saltier ocean near
the surface as a result of liquid brine rejected from the ice (A. K. Turner & Hunke, 2015;
Bailey et al., 2020; Singh et al., 2021; DuVivier et al., 2021).
4 Turbulence

Southern Ocean turbulence is driven by a wide range of processes and acts on many different scales. Turbulence is inherently characterised by nonlinear and chaotic motions. It is often difficult to establish the drivers of turbulence, which makes quantifying and categorising turbulence a challenge. Here, Southern Ocean turbulence is broadly categorised into eddies, jets and fronts (§4.1), convection (§4.2) and mixing (§4.3). Eddies, jets and fronts lie broadly in the realm of mesoscale turbulence, close to geostrophic and hydrostatic balance. Mesoscale turbulent processes are affected by a large range of factors, such as wind and buoyancy forcing, along with interactions with the mean flow, eddies, topography and more. Convection is driven by vertical buoyancy differences and is characterised by vigorous vertical motion and turbulent plumes. It can be confined to the upper ocean or extend to depth as polynya convection ( §§4.2.2–4.2.3). Mixing refers to three-dimensional turbulent processes that act to blend waters of different properties. To help categorise the wide range of processes that contribute to turbulence, we break the Southern Ocean into upper, interior and bottom layers (Fig. 16).

There exist past reviews and books on various aspects of ocean turbulence. For eddies, jets and fronts, the review by A. F. Thompson et al. (2018) (also mentioned in §2) considers the Antarctic Slope Current, an area of strong mesoscale turbulence processes, Ferrari and Wunsch (2009) discusses the energy framework for oceans, and McGillicuddy Jr (2016) examines a range of interactions at the oceanic mesoscale. For convection, J. Marshall and Schott (1999) review open ocean convection across the whole of the Earth’s oceans, while Morales Maqueda et al. (2004) (also mentioned in §3) reviews polynyas, including polynya convection and dense water formation. For more detailed reviews of mixing processes, the reader is referred to Whalen et al. (2020), Moum (2021) and Gille et al. (2022), as well as other relevant chapters of the recent Ocean Mixing book by Meredith and Naveira Garabato (2021).

4.1 Eddies, Jets and Fronts

The Southern Ocean is renowned for having one of the strongest turbulence fields in the global ocean, as measured by eddy kinetic energy (Ferrari & Wunsch, 2009). A common definition of eddy kinetic energy is the kinetic energy of deviations from the time-mean velocity field (A. R. Robinson, 1983). Most of this energy is found in the form of mesoscale turbulence, defined here as nonlinear motion close to geostrophic and hydrostatic balance. Mesoscale turbulence spreads energy across a broad range of length scales through nonlinear interactions, resulting in a complex, highly inhomogeneous and unsteady state of motion (Rhines, 1979). Because of the latitudinal dependence of the Coriolis parameter, the mesoscale range varies over the Southern Ocean, from 1–10 km near the Antarctic continent to 100–1000 km in the Antarctic Circumpolar Current (Fig. 17).

A generic feature of mesoscale turbulence is its tendency to form long-lasting, spatially localised features, such as jets (narrow, quasi-zonal currents), fronts (sharp, quasi-meridional gradients in temperature or salinity), and eddies (spatially and/or temporally coherent vortices). There is no uniquely accepted definition of eddies, jets or fronts (Chapman et al., 2020). For example, the dominant circulation feature of the Southern Ocean is the Antarctic Circumpolar Current (§2.1), which is composed of numerous jets that interact with each other (A. F. Thompson, 2008), flanked by sharp fronts, and co-located with the most active eddy field in the global ocean (Fu et al., 2016). The view is further complicated by the strong feedbacks that exist between these features. For example, jets become baroclinically and/or barotropically unstable to generate eddies, while eddies can flux momentum to sharpen jets (Waterman & Hoskins, 2013). Thus, this review tends towards aggregating these features into the broad category of mesoscale geostrophic turbulence.

In order to provide an overview of the dynamics of Southern Ocean mesoscale turbulence, we examine the sources, interactions and sinks in the eddy kinetic energy budget.
Figure 16. Schematic to illustrate surface, interior and bottom boundary layers in the Southern Ocean, with a summary of turbulence processes acting in each layer. The ocean colours indicate the density, from lighter (dark orange) to denser (dark blue) waters, and isopycnal contours are the interfaces between the layers. Note that the three layers are offset in latitude and disconnected in the vertical, with the surface layer 0–300 m depths, interior layer is 1000–4000 m and bottom layer 4500–5000 m. The water masses shown are Subantarctic Mode Water (SAMW), Antarctic Intermediate Water (AAIW), Circumpolar Deep Water (CDW), and Antarctic Bottom Water (AABW). Also shown on the top panel are the Antarctic Circumpolar Current (ACC), Polar Front (PF) and the Subantarctic Front (SAF).
Figure 17. Mesoscale turbulent structures are ubiquitous in the Southern Ocean. The Rossby number, defined as the vertical component of relative vorticity ($\partial v/\partial x - \partial u/\partial y$) divided by the planetary vorticity ($f$), highlights dynamical features throughout the Southern Ocean. The four inset figures show: the Antarctic Slope Current (top right), a large-scale meander near the Macquarie ridge and the associated energetic mesoscale eddy field (bottom right), the spatial variation in the dominant dynamical scale in the Southern Ocean (bottom left), and the highly energetic turbulence field in Drake Passage (top left). The velocity fields are from a snapshot from a regional simulation around Antarctica at a 1/20° lateral resolution, performed with the Modular Ocean Model, version 6 (Adcroft et al., 2019) by the Consortium for Ocean and Sea Ice Modelling in Australia (Kiss et al., 2020).

The primary source of eddy kinetic energy in the Southern Ocean is the generation of instabilities in the large-scale flow, ultimately powered by energy input from the wind (§4.1.1) and buoyancy forcing (§4.1.2). The equilibrium value of eddy kinetic energy in any region is governed by the energy source and redistribution of eddy kinetic energy by the background flow and other interactions, and also by the rate at which eddies dissipate their energy (§4.1.6). This view of the eddy kinetic energy reservoir as a source-sink problem makes it clear that a full understanding of the eddy field requires knowledge of both eddy formation processes and eddy dissipation dynamics.

This review will consider the following direct influences on the mesoscale turbulence field. The mesoscale turbulence field is influenced directly via exchanges of energy with
internal waves (§ 4.1.3). Feedbacks between different features can redistribute and influence energy via self-interaction of the mesoscale turbulence field (§ 4.1.4). Topography plays an important role in modulating the mesoscale dynamics of the Southern Ocean and connecting the large-scale circulation to smaller-scale, faster processes (§ 4.1.5). The geostrophic turbulence field is influenced indirectly by other components of the ocean–atmosphere–cryosphere system via their modulation of energy input by wind and buoyancy forcing. Thus, mesoscale turbulence acts as the bridge between the global-scale circulation and small-scale processes in the Southern Ocean.

4.1.1 Wind forcing

The power input from the atmosphere into the ocean is determined by the surface wind stress. The wind stress describes an exchange of energy and momentum between the air and the water, which is mediated through the sea surface and includes influences from surface gravity waves (§ 5.1). Wind stress is often calculated via a bulk formula, which implies that it is proportional to \(|U_{\text{air}} - u| (U_{\text{air}} - u)|\), where \(u\) is the ocean surface velocity and \(U_{\text{air}}\) is the wind velocity at a reference height (typically 10 m) above the sea surface. Two features in the wind stress bulk formula are noteworthy. First, the wind stress is quadratic in velocity, which implies that the fluctuating winds need not average out but can contribute towards total wind stress felt by the ocean. Second, the wind stress depends on the relative flow between the atmosphere and the ocean, \(U_{\text{air}} - u\), and therefore the ocean flow affects how the ocean feels the atmosphere.

For a long time it was thought that most of the wind energy input resulted from the correlation between the mean wind stress and mean currents, and that the time-varying wind and ocean flow variability contribution was negligible (Wunsch, 1998; Scott & Xu, 2009). However, more recent studies have highlighted the important role of the synoptically varying winds (here, this refers to winds varying on “short” timescales of hours to days), which can result in a 70% increase in power input into the ocean from the winds (Zhai et al., 2012). Most of this energy enters the ocean in the winter time and in regions with strong synoptic wind variability, such as the Southern Ocean (Flexas et al., 2019). The wind stress injects energy into both geostrophic and higher frequency motions (especially near the inertial frequency) and from the latter, near-inertial waves are energised that propagate down below the mixed layer into the deep ocean (§ 5.3).

Generally, ocean velocities are much smaller than wind velocities, and, therefore, one might expect that the relative flow contribution to the wind stress power input would be insignificant. However, ocean flow features appear in much smaller length scales and vary at much longer time scales than the synoptic variability of the winds. If the relative wind and ocean flows are opposing, then winds damp the ocean flow and remove energy from the ocean, particularly in eddy-rich regions like the Southern Ocean (Zhai et al., 2012). The relative wind effect has a particularly large impact on mesoscale turbulence through “eddy killing”, which results in a 20–40% reduction in mesoscale eddy kinetic energy compared to a formulation of the surface stress that does not take ocean currents into account (e.g., Renault et al., 2016; Jullien et al., 2020).

The presence of sea ice alters the relationship between atmospheric winds and wind stress. In regions with drift sea ice (§ 3.2.1), the momentum transfer can be three times that in an ice free ocean (T. Martin et al., 2014). However, at higher concentrations, the internal stresses in sea ice can reduce the momentum transfer into the ocean, potentially even resulting in an ice-ocean drag that decelerates ocean currents (Meneghello et al., 2018; A. L. Stewart et al., 2019). Landfast sea ice and ice shelves are critical elements of the coastal cryosphere through their complete removal of wind stress forcing (§ 3).
4.1.2 Buoyancy forcing

Buoyancy forcing is another driver of mesoscale geostrophic turbulence in the Southern Ocean. The presence of stratification allows baroclinic modes of instability to generate geostrophic turbulence, while also weakening the barotropic potential vorticity constraints on geostrophic flow (Cushman-Roisin & Beckers, 2011). The large scale meridional sloping of isopycnals across the Antarctic Circumpolar Current region is maintained by the wind (Ferrari & Wunsch, 2009). Mesoscale turbulence is tightly coupled to stratification by working to flatten these isopycnals. The resulting stratification with flattened isopycnals affects the dynamics of the Southern Ocean at all scales. For example, increased heat storage north of the Subantarctic Front has been linked to an acceleration of the zonal flow (Shi et al., 2021). In addition, baroclinic instability is central to the dynamics of standing meanders of the Antarctic Circumpolar Current (Watts et al., 2016; Foppert et al., 2017; Youngs et al., 2017; Constantinou & Hogg, 2019). It has also been found that interactions between Southern Ocean jets, topography, and stratification can lead to rapid changes in ocean ventilation (Klocker, 2018).

Southern Ocean stratification is influenced by many processes, which can also go on to impact mesoscale turbulence. Some processes, such as sea ice melt and surface heating, act to stratify the water column (Haumann et al., 2020). Others, such as convection and mixing, decrease the vertical stratification. Meltwater from ice sheets and ice shelves leads to fresh, cold surface water near Antarctica. For example, the meltwater plume from ice shelf melting modifies the ocean stratification and uptake of surface buoyancy, which will go on to influence the mesoscale turbulence field. Fast ice reduces ocean–atmosphere heat and salt exchanges, replacing them with ice–ocean exchanges of melting and freezing. Vertical mixing by mesoscale turbulence underneath sea ice dissipates eddy kinetic energy and reduces sea ice thickness by up to 10% (Gupta et al., 2020). Strong horizontal density gradients from vertical convective mixing can provide energy for geostrophic turbulence to restratify that region (Jones & Marshall, 1997; Kurtakoti et al., 2018).

4.1.3 Internal wave interactions

The geostrophic turbulence field in the ocean continuously exchanges energy with the internal wave field (E. D. Brown & Owens, 1981; Polzin, 2010; Polzin & Lvov, 2011). Internal waves “see” eddies and jets as a slowly moving and usually larger-scale flow from which they can both extract or input energy, depending on the relative direction of the eddying flow and wave propagation. It has been argued that energy exchange with internal waves is a significant net sink of eddy energy (Polzin, 2008, 2010), although other studies in the Southern Ocean have found the opposite effect (Cusack et al., 2020; Shakespeare & Hogg, 2019). As such, the overall effect of internal waves on eddies and jets remains a topic of active research (§ 5.3).

4.1.4 Mesoscale turbulence self-interactions

Mesoscale turbulence in the Southern Ocean exhibits many of the nonlinear self-interactions seen in two-dimensional and quasi-geostrophic turbulence under the constraints of rotation and stratification (Hopfinger & Van Heijst, 1993). The level of eddy self-interaction can be quantified using a nonlinearity parameter, which expresses the ratio of the rotational velocity of the eddy to its translational velocity (Chelton et al., 2011; Klocker et al., 2016). Southern Ocean eddies, particularly in the Antarctic Circumpolar Current, typically have large values of this parameter (of order 10), implying that the eddies cannot be regarded as linear perturbations to a quiescent background, but instead modify the surrounding flow by trapping and transporting fluid (Chelton et al., 2011). These self-interactions include eddy merging and splitting events (Cui et al., 2019), the formation of quasi-stable dipoles, quadrupoles and larger eddy assemblages (e.g., Gallet & Ferrari, 2020), and the cascade of energy from small to large scales (Salmon, 1998; Scott & Wang, 2005; Aluie et al., 2018;
The inverse energy cascade is consistent with a pronounced seasonal cycle in eddy kinetic energy and eddy diameter observed in a 25-year climatology of satellite altimetry measurements (Martínez-Moreno et al., 2021), where small-scale (diameter < 120 km) coherent eddies peaked in amplitude in mid-summer, while large-scale (> 120 km) eddies peaked in autumn. The findings suggest an inverse cascade from small scales, driven by baroclinic instability early in the summer, to large diameter eddies which grow in amplitude later in the season.

Eddy-mean flow interactions are mediated by eddy fluxes of buoyancy and momentum (Q. Li et al., 2016). For example, strong jets become baroclinically and/or barotropically unstable to generate eddies (H. E. Phillips & Rintoul, 2000; Chapman et al., 2015; Watts et al., 2016; Youngs et al., 2017; Foppert, 2019; Constantinou & Hogg, 2019), while eddies can flux momentum to sharpen jets (Waterman & Hoskins, 2013). Eddy momentum fluxes act to accelerate (for a converging momentum flux) or decelerate (for a diverging momentum flux) the Antarctic Circumpolar Current near topographic features such as the Drake Passage and Campbell Plateau (Morrow et al., 1994; Ivchenko et al., 1997; R. G. Williams et al., 2007). Eddy buoyancy fluxes play an important role in setting the slope of isopycnal surfaces, thereby influencing the strength and stability of the Antarctic Circumpolar Current (Karsten et al., 2002; J. Marshall & Radko, 2003; Olbers et al., 2004; Olbers & Visbeck, 2005).

Eddy geometry (the shape, size, and anisotropy of eddies) provides a useful framework for characterising eddy-mean flow interactions (D. P. Marshall et al., 2012; Waterman & Lilly, 2015). Two distinct interactions between eddies and the time-mean jet have been identified in an idealized quasigeostrophic jet system (Waterman & Hoskins, 2013). Firstly, eddy-mean flow interactions arise from eddy effects that stabilize the jet to barotropic and baroclinic instability through down-gradient fluxes of potential vorticity. Secondly, eddy-mean flow interactions drive time-mean recirculations via an upgradient potential vorticity flux driven by eddy enstrophy convergence (Waterman & Jayne, 2012).

4.1.5 Topographic effects

Bottom topography plays an important role in modulating the Southern Ocean mesoscale turbulence field. Both models and observations indicate that enhanced eddy kinetic energy, cross-frontal transport, and eddy-induced upwelling can be found downstream of major topographic features (Fig. 18; Foppert et al., 2017; Tamsitt et al., 2018; Barthel et al., 2022; Yung et al., 2022). Topography also plays a pivotal role in modulating jet evolution. Observations and idealised models show that the formation of jets, their meridional spacing and variability, and associated transport depend on the length scale and steepness of the topographic features (A. F. Thompson, 2010; Boland et al., 2012; Chapman & Morrow, 2014; Freeman et al., 2016; Constantinou & Young, 2017).

Near the Antarctic margins, ice topography can also influence mesoscale geostrophic turbulence. Rapid changes in water column thickness near ice shelf and glacier tongues modify local angular momentum balances (van Heijst, 1987). Strong vorticity gradients occur at ice-shelf fronts (Steiger et al., 2022), where the ice draft may be greater than half the local seabed depth. The ice creates a barrier against which water may pool and a strong along-front flow may develop (Malyarenko et al., 2019). There is evidence that this front provides an impediment to barotropic processes but that baroclinic transport can persist (Wählin et al., 2020; Steiger et al., 2022) enabling penetration of heat beneath ice shelf frontal regions (C. L. Stewart et al., 2019; Davis et al., 2022).

4.1.6 Dissipation of eddy kinetic energy

The primary source of eddy energy is the generation of instabilities in the large-scale flow, ultimately powered by wind stress and buoyancy forcing (§§ 4.1.1–4.1.2). In the Southern Ocean, both barotropic and baroclinic instability contribute to the eddy field, although
baroclinic instability is expected to dominate at the mesoscale (Youngs et al., 2017). However, the mesoscale energy is known to have a largely upscale cascade, meaning that energy is returned to the large-scale flow field. This upscale cascade is a consequence of the conservation of potential vorticity and theoretically applies to balanced flows at low Rossby number and in the interior of the ocean (Rhines, 1977). It follows that situations in which balance is broken yield the possibility of a forward cascade of energy to the submesoscales, internal waves and shear-driven turbulence. Questions arise regarding the mechanisms that transfer energy to smaller scales (i.e., the direct cascade) down to length scales where viscous processes can act to remove energy from the flow.

The main candidate mechanisms for loss of balance involve interactions at the boundary of the ocean, either the surface or bottom. At the surface, eddies can generate filaments leading to “frontogenesis”, thereby breaking the constraint of low Rossby number flow and creating an active submesoscale flow field (Barkan et al., 2015; McWilliams, 2021). Submesoscale instabilities near sloping bottom boundaries (Callies, 2018; Wenegrat et al., 2018; Wenegrat & Thomas, 2020) may also drive loss of balance, although research in this area is still in its infancy. Zhai et al. (2010) proposed that western boundary currents act as an “eddy graveyard”, supported by satellite measurements. Such a mechanism likely involves interactions between eddies and shoaling topography, such as frictional (Evans et al., 2020; Wright et al., 2013) or dynamical (Dewar & Hogg, 2010) mechanisms. Additionally, the rotation of coherent vortices leads to a wind-stress torque that directly damps eddies (Zhai et al., 2012).
One dynamical mechanism that removes energy from eddies is the generation of internal waves from eddy-topography interaction (§ 5.3.1). The intense and deep reaching mesoscale flow of the Southern Ocean results in bathymetric interactions that generate Doppler-shifted internal waves, the most well-known of which are lee waves. The breaking of these waves (§ 5.1.2) exerts a drag on the background mesoscale flow. A. C. Naveira Garabato et al. (2013) evaluated time-mean lee wave drag globally and found that, while it is a minor contributor to the ocean dynamical balance over much of the ocean, it is a significant player for Antarctic Circumpolar Current dynamics in the Southern Ocean. Extending this estimate to transient eddies in the Southern Ocean, Yang et al. (2018, 2021) have shown that lee wave drag processes dominate over the turbulent bottom boundary layer drag for eddy dissipation, a result consistent with previous results from higher resolution idealised models (Nikurashin et al., 2013).

It has also been proposed that loss of balance can occur spontaneously in the ocean, in the absence of surface forcing or bottom interactions (Molemaker et al., 2005; Shakespeare, 2019). Simulations show that spontaneous emissions of internal gravity waves occurs in balanced flow, but that the energy transferred in this process, while it may be locally important, is unlikely to be a regionally or globally significant sink of eddy energy (Vanneste, 2013; Nagai et al., 2015; Shakespeare & Hogg, 2017; Chouksey et al., 2018). Alternatively, the exchange of energy between eddies and surface- or bottom-generated internal waves can, in some circumstances, result in a net extraction of energy from the eddy field into internal waves (§ 5.3.2).

Despite the range of available mechanisms for eddy dissipation, there is no clear view of which mechanism dominates, nor a demonstration of the relative magnitude of these mechanisms in the Southern Ocean.

4.2 Convection

Convection is a type of flow driven by a vertical buoyancy differential that, in the ocean, is due to temperature and salinity differences. Unstable buoyancy differences, such as cold and/or saline water overlying warm and/or fresh water, trigger small-scale three-dimensional motions commonly known as “turbulent convection”. Turbulent convective events are often characterised by plumes that vertically flux buoyancy and mix with the ambient ocean. Buoyancy loss through various surface drivers (net cooling, evaporation, and sea ice formation) is a primary mechanism for triggering this turbulent convection and dense water formation. The dominant convection processes influencing Southern Ocean dynamics are mixed layer convection and polynya (coastal and open ocean) convection, as discussed in the following sections.

4.2.1 Upper mixed layer convection

Turbulent convection in the upper ocean occurs when surface cooling, evaporation and/or brine rejection leads to a gravitationally unstable water column. Various surface forcings (e.g., wind stress, heating and cooling, evaporation and precipitation, freshening and brine rejection during sea ice melting and formation) drive small-scale eddies that trigger turbulent convection. Turbulent convection is strongly linked to the depth of the mixed layer, which is the uppermost part of the ocean characterised by a homogeneous density distribution. Over the Southern Ocean, the mixed layer experiences a strong seasonal cycle and is deeper during the Austral winter and shallower during the Austral summer (Fig. 19; Sallée et al., 2006; Dong et al., 2008; Ren et al., 2011; Pellichero et al., 2017; Buongiorno Nardelli et al., 2017). In broad terms, the deep winter mixed layer is mostly driven by convective processes, either from temperature inversions during surface cooling or salinity inversions during brine rejection, or from a combination of these two effects (Pellichero et al., 2017; Clément et al., 2022). Convection becomes less pronounced during summer due to the increased stability in the water column from surface heating and sea ice melting.
Most Southern Ocean regions experience moderate to strong seasonality resulting in a large variation of heat and salt fluxes at the ocean surface. There are seasonal variations in the air–sea fluxes that result in changes in heating and cooling at the ocean surface. There are also strong seasonal variations of sea ice–ocean fluxes due to the large extent of sea ice cover and its associated seasonal changes. Mixed layer properties and dynamics are very different between sea ice covered and sea ice free zones due to different characteristics of convection and related water mass transformation. The spatial variation of the mixed layer depth is more pronounced in the latitudinal direction due to both variation of air–sea and ice–ocean fluxes, leading to a meridional banded structure of the winter mixed layer depth across the Southern Ocean. This band is deep near the Antarctic continent, becoming shallower farther offshore, before deepening again along the northern flank of the Antarctic Circumpolar Current (Fig. 19c,f Pellichero et al., 2017; Wilson et al., 2019). The winter deep mixed layer region to the north of the Subantarctic Front is where Subantarctic Mode Water is formed (McCartney, 1977).

In the region free of sea ice, the seasonal cycle of air–sea interactions plays a dominant role in the seasonal variation of the heat content in the mixed layer (Sallée et al., 2006; Dong et al., 2007, 2008; Pellichero et al., 2017) with warming of the subsurface ocean during spring and summer and cooling during autumn and winter. A large buoyancy loss from the ocean to the atmosphere during wintertime causes an unstable temperature inversion leading to vertical convection, which results in a deep mixed layer. Thus, net vertical heat flux at the ocean surface (negative during most autumn and winter months spanning March to September) is the dominant term in the heat budget of the mixed layer (Pellichero et al.,
The second largest term in the heat budget is vertical entrainment of the cold ambient water, which is negative on average and cools the mixed layer (Dong et al., 2007). Horizontal Ekman advection of cold water from the south due to strong winds across the Southern Ocean also plays some role in cooling the upper ocean throughout the year. A density inversion in the upper ocean is observed over a wide area spanning the Antarctic Circumpolar Current and further north, including the mode water formation regions.

Sea ice covers a major part of the Southern Ocean in winter (Fig. 9), insulating the ocean from the cold atmospheric air and minimising the heat loss. The start of winter sees sea ice formation resulting in brine rejection and cold surface waters, which leads to a top-heavy water column susceptible to convective instabilities. The stability of the water column is determined by both the stability of the temperature profile and the salinity profile (Pellichero et al., 2017). The heat and salinity fluxes at the sea ice interface are the dominant contributors to the respective heat and salinity budgets, with negligible contributions from lateral advection (by Ekman transport) and diffusion. The entrainment of salt flux from the bottom of the mixed layer plays an important role in the salinity budget of the mixed layer. From late autumn onward, the deepening of the mixed layer entrains the underlying, relatively salty Circumpolar Deep Water into the mixed layer in the Weddell Sea and Ross Ice Shelf regions, decreasing the overall buoyancy content of the mixed layer. The degree to which the Circumpolar Deep Water interacts with the mixed layer varies around the Antarctic continent. For example, in the East Antarctic, the strong Antarctic Slope Current and easterly winds tend to inhibit the entrainment of the Circumpolar Deep Water into the surface mixed layer (A. F. Thompson et al., 2018). In addition to the above processes, leads exist in many sea ice covered areas there (§3.2.2; Muchow et al., 2021), which allow for the direct interaction between the cold atmosphere (frequently below −30°C) and the ocean, forming large localised convection that is dominantly driven by sensible heat loss and brine rejection (S. D. Smith et al., 1990; Simmonds & Budd, 1991).

In early to mid-winter the heat flux from the ocean, which warms the sea ice, is much less than the heat loss to the atmosphere through the upper surface of the ice, resulting in rapid sea ice growth and both temperature and salinity driven convection (Wilson et al., 2019). As the under-ice mixed layer deepens, it cools to about freezing point while also becoming saltier. The deepening of the mixed layer is associated with an upward heat flux of 20–60 W m$^{-2}$, increasing the mixed layer temperature to 0.05–0.2°C (Gordon & Huber, 1990; Keeling & Visbeck, 2011). This entrainment provides an efficient mode for exchanging freshwater along with heat and atmospheric gases (e.g., carbon dioxide, oxygen) between the deep ocean and the atmosphere (Gordon, 1991). The entrainment of warm water continues to increase the heat flux from ocean to ice. In late winter, when the ocean heat flux to the sea ice is equal to or more than the heat loss to the atmosphere, the entrained heat melts the sea ice from below, and a strong surface stratification establishes due to the release of freshwater from melting that can rapidly slow down surface-driven convection and mixed layer growth. However, turbulence can also be sustained by double-diffusive convection processes as cold and freshwater overlie warm and salty water (§3.1.6). Evidence for double-diffusive convection was reported in observations of the subsurface water column both in the Weddell and Ross Seas during late winter time (Shaw & Stanton, 2014; Bebieva & Speer, 2019).

4.2.2 Coastal polynya convection

The ocean around Antarctica is covered in sea ice during much of the year, particularly in winter, except for pockets of open water known as polynyas (Morales Maqueda et al., 2004). Polynyas generally lie close to the coast, with strong katabatic winds blowing any newly-formed sea ice away from the coast. Coastal polynyas are key regions for water mass transformation via atmosphere–sea ice–ocean interactions (Killworth, 1983; Tamura et al., 2008). The process of coastal polynya convection begins at the surface, where there is buoyancy loss due to a sudden cooling or an increase in sea ice production and brine rejection,
or a combination of both of these effects. In some circumstances, polynya convection is started by brine rejection and then maintained by surface cooling, as convection continues to bring warmer waters to the surface.

Buoyancy loss, from brine rejection or surface cooling, causes deepening of the upper ocean mixed layer followed by convection that can reach the ocean floor (J. Marshall & Schott, 1999). The ocean floor on the Antarctic continental shelf is typically a few hundred metres deep, extending to 1 km near the shelf break (Amblas & Dowdeswell, 2018). The convection region or “patch” is made up of convective plumes of around 1 km or less in width. Baroclinic eddies form from the edge of the convective patch, due to the strong horizontal gradient in buoyancy between the convective region (less buoyant) and surrounding ocean (more buoyant). However, these eddies may be dissipated by the neighbouring ice shelf or, farther from the coast, the sea ice cover. The width of these eddies will depend on the Rossby radius of deformation which is roughly 5–10 km in coastal polynya regions (e.g., the Rossby radius is approximately 4 km near the Ronne Ice Shelf; Arthun et al., 2013). Sustained coastal convection is dependent on a number of driving factors coalescing under the right conditions. In particular, coastal polynya regions need continual access to the warm, salty Circumpolar Deep Water heat reservoir that drives heat loss to the atmosphere and rapid sea ice melt. Surface winds (katabatics and easterlies) are required to promote favourable conditions for sea ice formation and continued northward export of sea ice.

Surface water mass transformation in polynyas is often seasonal and localised. While some polynyas are strong factories of convection and dense water formation throughout large portions of the year, other polynyas do not produce significant dense water mass. Some of the most productive regions of dense water formation are the Weddell Sea, Prydz Bay, Adélie Land and Ross Sea (Morales Maqueda et al., 2004). Polynya convection can also undergo changes if the surface conditions change from year-to-year. Large grounded icebergs can act, essentially, as islands, leading to modified convection and ocean circulation. For example, a polynya in Adélie Land was noted to decrease in dense water formation (leading into Adélie Land Bottom Water) after newly-formed sea ice was blocked from exiting the polynya region by a large grounded iceberg (Snow et al., 2018). Observations and modelling also demonstrate that meltwater plumes from neighbouring ice shelves may freshen the surface waters in the polynya region and result in a shut down of convection (Silvano et al., 2018; Moorman et al., 2020). As polynya convection can remove heat from Circumpolar Deep Water as it comes onto the continental shelf, a decrease in polynya convection can result in unmodified Circumpolar Deep Water reaching the base of neighbouring ice shelves, exacerbating ice shelf melt. This can then have a feedback effect on the convection in polynyas downstream, resulting in further reductions in convection (Silvano et al., 2018).

4.2.3 Open ocean polynya convection

Open ocean convection is characterised by the rapid vertical heat exchange between the surface and deep ocean, driven predominantly by sensible heat loss or brine rejection at the surface of the ocean, and relatively unencumbered by local coastal processes (J. Marshall & Schott, 1999). It occurs further offshore than coastal convection and is a more intermittent phenomenon. In regions where open ocean convection is active, gaps in the sea ice cover (polynyas) emerge and can persist for weeks and up to several months (Comiso & Gordon, 1987). Such polynyas have been observed in the Weddell Sea in 1974 (Gordon, 1978), and also to a lesser extent in 2016 and 2017 (Jena et al., 2019), and in the western Cosmonaut Sea (persistent in Austral autumn and winter; Comiso & Gordon, 1987).

The life cycle of a typical open ocean convection event is relatively well-understood (J. Marshall & Schott, 1999). In the first preconditioning phase, favourable local oceanic conditions are set up that lower the thermodynamic barrier to rapid sensible heat exchange with the atmosphere (Fig. 20a). In the second deep convection phase, deep, turbulent ocean convection is triggered which spawns multi-scale convective chimneys and a geostrophic rim
Figure 20. Different stages of open-ocean convection shown in high-resolution direct numerical simulations. Figure reproduced from Vreugdenhil and Gayen (2021) and based on the simulations by Sohail et al. (2020).

current (Fig. 20b). Finally, given the right conditions, the rim current becomes baroclinically unstable, pinching off high-buoyancy baroclinic eddies which rapidly laterally mix the convective patch in the third lateral spreading phase (Fig. 20c). If favourable forcing conditions persist, the convective event will reach a quasi-equilibrium state in the lateral spreading phase, with minimal changes to the mixed layer or net vertical heat flux. The convection will only cease when the subsurface heat reservoir has been depleted, or freshwater input at the surface occurs, acting to restabilise the water column. Once such conditions cease, baroclinic eddies rapidly break down the convecting patch via lateral mixing, restratifying the ocean and encouraging reformation of sea ice (Jones & Marshall, 1997).

In the Southern Ocean, the Weddell Sea is a critical region for open ocean polynya formation and convection. In the Weddell Sea, Maud Rise polynyas have been intermittently observed around the Maud Rise seamount region, with the most recent notable example being in 2016 and 2017 (Fig. 21). In the mid-1970s, such Maud Rise polynyas were a precursor to the much larger and more consequential Weddell polynya, which emerged in 1974 and persisted through to 1976 (Fig. 21). Over its life, the Weddell polynya reached a maximum extent of 250,000 km$^2$, generated dense water at an average rate of 1.6–3.2 Sv, and reduced the heat content of the underlying Weddell Deep Water by $12.6 \times 10^{20}$ J (§ 2.5 Gordon, 1982).

Given the dearth of direct observations in the Weddell Sea region, much research has focused on the drivers that precondition Maud Rise polynyas and yield larger Weddell polynyas in the Southern Ocean. Several candidate processes have emerged that, through complex interactions, likely dictate the emergence and strength of the Maud Rise polynya and Weddell polynya. First, a period of prolonged negative Southern Annular Mode conditions, aided by La Niña, can salinify the ocean surface and reduce the stability of the water column (Gordon et al., 2007). Interactions of background flow with the Maud Rise (Fig. 21), particularly in these weakly stratified conditions, may give rise to a Taylor column (a stagnant region that can form over an obstacle in a rotating flow; G. I. Taylor, 1923) that is isolated to the seamount, bringing warm, salty Weddell Deep Water closer to the ocean mixed layer (Kurtakoti et al., 2018; Steur et al., 2007). Following this preconditioning, a negative wind stress curl over the Weddell Sea would strengthen the Weddell Gyre, causing the underlying Weddell Deep Water to upwell (Cheon et al., 2015) and melt sea ice in the region. Cyclonic eddies may also shed off the Maud Rise seamount, opening gaps in the sea ice and enabling rapid heat loss to the atmosphere (D. M. Holland, 2001). Intermittent cyclones can also provide a strong mechanical forcing, opening the sea ice pack and expos-
Once a Maud Rise polynya is triggered, westward propagation of the polynya can yield a larger Weddell polynya, especially if the heat reservoir in the Weddell Deep Water is large and the wind stress curl and Southern Annular Mode are strongly negative (Fig. 21; Kurtakoti et al., 2021). Note that Weddell polynya formation is not guaranteed once an Maud Rise polynya is formed. For example, the relatively large Maud Rise polynya in 2017 (Fig. 21) did not transition to a Weddell polynya, as a positive Southern Annular Mode index that year meant the water column was more stable and inhibited Weddell polynya formation (Cheon & Gordon, 2019). A Maud Rise polynya, or Weddell polynya, will persist in quasi-equilibrium (as described above) until it is destroyed by the loss of sub-surface heat, the input of surface freshwater, or through interactions with broader-scale gyre currents (D. M. Holland, 2001; Martinson et al., 1981).

4.3 Mixing

Three-dimensional turbulence and mixing in the Southern Ocean, whether in the interior or in the surface and bottom boundary layers, plays an important role in shaping air–sea and ice–ocean exchange (e.g., Holte et al., 2012; Rintoul, 2018), watermass transformation (e.g., Downes et al., 2011; Cerovecki & Mazloff, 2016; Evans et al., 2018) and tracer...
transport (e.g., Mashayek, Ferrari, et al., 2017; Uchida et al., 2020). Three-dimensional turbulence lies at the bottom of the spatial and temporal scale range, acting to absorb the down-scale cascade of energy and scalar variance generated by motions at larger scales, and ultimately remove it at molecular scales. The millimetre to centimetre scales of turbulence, coupled with its highly intermittent nature, make it extraordinarily difficult to measure. Thus, much of our knowledge on the distribution of mixing in the ocean is inferred from observations of larger scales.

The term “mixing” refers to the process of blending waters of different properties. The focus of § 4.3 is on the irreversible mixing of scalars. Diapycnal mixing or mixing across density surfaces is quantified using a diapycnal diffusivity, which is typically seven orders of magnitude smaller than the horizontal components that are set by along-isopycnal mesoscale stirring (de Lavergne et al., 2022). Mixing along isopycnals can facilitate diapycnal mixing by creating fine-scale gradients, for example of temperature, which are more readily acted upon by turbulence (Abernathey et al., 2022; de Lavergne et al., 2022). In addition, isopycnal mixing can lead to a densification via cabling, where mixing two water parcels of the same density results in a denser parcel due to nonlinearities in the equation of state (Urakawa & Hasumi, 2012; L. N. Thomas & Shakespeare, 2015; Groeskamp et al., 2016).

Direct measurements of mixing, resolving millimetre to centimetre scales, are limited to specialised research campaigns involving tracer release experiments (Ledwell et al., 2011) and microstructure instruments (Waterman et al., 2013; Laurent et al., 2012; Ferris et al., 2022; Fer et al., 2016). Microstructure instruments rely on rapid-response velocity, temperature or salinity sensors that resolve variations with depth on the scale of centimetres, and provide an estimate of the dissipation of turbulent kinetic energy $\varepsilon$ (or tracer variance). Diapycnal diffusivity is then estimated as $\Gamma \varepsilon/N^2$ (Osborn, 1980), where $\Gamma$ is generally assumed equal to 0.2 (Gregg et al., 2018) and $N^2$ is the square of the buoyancy frequency, defining the vertical stratification. These campaigns often focus on regions that are of interest from a geophysical fluid dynamics perspective and are not necessarily representative of the broader ocean.

Due to the limitations of direct observations throughout the Southern Ocean, finescale parameterizations of turbulent dissipation are widely used. Finescale methods applied to measurements of density and velocity that resolve the vertical length scales of internal waves can be used to infer the mixing that results from the breaking of internal waves (§5.3.2), either locally or after propagating some distance (Polzin, Naveira Garabato, Huussen, et al., 2014). The finescale method has two major assumptions: 1) all the observed shear and strain in the ocean interior is due to internal waves, and 2) the nonlinear interactions between the waves result in a downscale energy cascade leading to wave breaking and turbulence (Polzin, Naveira Garabato, Huussen, et al., 2014; Whalen et al., 2015). Fewer assumptions are required when both velocity and density profiles are measured simultaneously, again limiting the observations to research vessels (Waterhouse et al., 2014) and autonomous instruments that measure coincident profiles of velocity and density, such as Electromagnetic Autonomous Profiling Explorer (EM-APEX) profiling floats (Meyer, Sloyan, et al., 2015; Cyriac, Phillips, Bindoff, & Polzin, 2022).

The most broadly available estimates of mixing come from the global Argo profiling float array (Whalen et al., 2012, 2015) that measure profiles of temperature and salinity to 2000 m. The absence of ocean velocity profiles in these measurements requires an assumption of the ratio of shear variance to strain variance, often chosen between three and seven (Kunze et al., 2006; Cyriac, Phillips, Bindoff, & Polzin, 2022; Waterhouse et al., 2018). Parameterised estimates of mixing have been found to agree with direct measurements within a factor of two to three in the open-ocean thermocline (Whalen et al., 2015, 2020). This range of mixing observations provides some knowledge of the global-scale distribution of mixing and its seasonal variability, which has been shown to be closely correlated with seasonal variations in wind strength. However, in the Southern Ocean, apart from targeted field campaigns, there is little knowledge of the amplitude and variability of mixing in the
surface mixed layer, below 2000 m depth, in boundary currents, in ice-covered regions, and at spatial scales smaller than 100 km and temporal scales less than a month.

We organise §4.3 by separately considering mixing within the surface boundary layer (§4.3.1), the interior (§4.3.2) and near the bottom (§4.3.3). Fig. 16 illustrates schematically the three layers and summarises the processes affecting mixing that will be addressed in the following sections.

### 4.3.1 Upper ocean mixing

Air–sea exchanges in the Southern Ocean are mediated through the surface mixed layer and, thus, are shaped by boundary layer mixing. Surface boundary layer mixing is fundamental to surface ventilation and hence water mass formation (§§ 2, 4.2). Fox-Kemper et al. (2022) provides a thorough review of surface boundary layer mixing processes. The depth of the surface boundary layer is also important to the input of wind power that drives near-inertial oscillations and internal waves that ultimately contribute to deeper ocean mixing (§5.3). The Southern Ocean surface is characterized by strong time-mean and time-variable wind stress, large lateral density gradients and strong seasonally-varying heat and freshwater fluxes. The resulting transient near-surface mixing geography is shaped by a myriad of processes including surface waves (Belcher et al., 2012; Fox-Kemper et al., 2022), submesoscale and frontal dynamics (Du Plessis et al., 2019; Giddy et al., 2021; Gula et al., 2022), wind-generated near-inertial waves (Whitt et al., 2019; Whalen et al., 2020), which also penetrate into the interior to influence interior mixing (Alford et al., 2012; Cyriac, Phillips, Bindoff, & Feng, 2022), and sea ice interactions (Pellichero et al., 2017; Evans et al., 2018; S. Swart et al., 2020). Further, recent work highlights the interaction between mixing, air–sea heat fluxes and sea ice formation, leading to the transformation of Circumpolar Deep Water and the formation of intermediate water (§2; Evans et al., 2018). While an understanding of these processes is developing, observations are sparse and parameterization development has so far been based on Northern Hemisphere data.

Surface gravity waves play a vital role in both air–sea exchange and deepening of the surface mixed layer through entrainment (Fig. 24, §5.1.2). The bubbles, spray and foam resulting from breaking surface waves lead to a complex multiphase fluid that is a challenge to both observe and model. This multiphase fluid is critical to both air–sea fluxes and can also affect surface roughness and wave dynamics. Surface waves contribute to mixed layer entrainment through the formation of deeply penetrating Langmuir turbulence and non-breaking wave turbulence. Langmuir cells are driven by the interaction between the wind-driven shear current and the Stokes drift current and result in pairs of parallel counter-rotating vortices oriented in the downwind direction. Belcher et al. (2012) concluded surface wave-forced Langmuir turbulence should be a major source of turbulent kinetic energy in the Southern Ocean. Langmuir cells can contribute to entrainment even when the cells do not reach the mixed layer base by enhancing the shear via pressure work (Q. Li & Fox-Kemper, 2020). Non-breaking (irrotational) surface waves can enhance existing background ocean turbulence when the orbital velocities of the irrotational waves interact with them (Qiao et al., 2016). Observations showed that they have capacity to deepen the mixed layer depth (Toffoli et al., 2012). Due to the extreme wave environment of the Southern Ocean, it is likely that these processes play a key role. However, observations remain sparse due to its remoteness, harsh environment, and sea ice cover. Therefore, ocean circulation models have been used to gain insight into these processes. Simulations of the surface boundary layer at the West Antarctic Peninsula that include parameterization of Langmuir cells demonstrate more realistic deep mixed layers on the slope and shelf regions due to Langmuir entrainment (Schultz et al., 2020). However, the first extensive microstructure turbulence observations of the Southern Ocean surface boundary layer show that Langmuir circulations alone do not explain the enhanced turbulence at the base of the mixed layer. Instead, storm forced inertial currents provide additional shear (Ferris et al., 2022).
Large inertial oscillations can be resonantly excited in the mixed layer when strong winds turn with the inertial rotation (Dohan & Davis, 2011). These inertial oscillations can then leave the mixed layer as propagating near-inertial waves. The inertial waves induce shear within the base of the mixed layer in the so called “mixing transition” layer, which results in mixing and widening of the layer (Skyllingstad et al., 2000; Forryan et al., 2015). High-resolution turbulence observations and drifter data show that the inertial oscillation-induced turbulent dissipation rate across the layer is an order of magnitude larger than that induced by most other mixed layer processes (with the exception of mixed layer frontal instabilities), thereby further highlighting the importance of wind-driven inertial oscillations for thermocline mixing (Peng et al., 2021). At first glance, Southern Ocean density profiles appear to have much deeper surface mixed-layers (hundreds of metres) than is typical in more temperate regions. However, this layer is actually weakly but stably stratified, with the active mixing layer confined close to the surface (Kilbourne & Girton, 2015). Therefore, a “slab model” (an analytical model that treats the surface mixed layer as a slab to estimate the mixed layer response to wind stress) can be applied to the actively mixing layer to estimate the near-inertial response to wind input (Pollard & Millard Jr, 1970). Southern Ocean observations demonstrate that near-inertial internal waves are responsible for transporting large amounts of energy downward to the base of the mixing layer where (indirect) estimates of vertical diffusivities are found to be enhanced (Ferreira Azevedo et al., 2022).

The upper ocean mixing is also impacted by complex, horizontal processes emanating from fronts, eddies and jets occurring at small spatial scales that extend down to the submesoscale (tens of centimetres to tens of kilometres, and hours to days). In the Southern Ocean, the strong surface forcing, persistent lateral density gradients, weak vertical stratification and deep mixed layers further enhance submesoscale mixing (Gille et al., 2022). Submesoscale instabilities, induced by the large-scale adiabatic mesoscale stirring (§ 4.1), can lead to strong subduction of water (K. A. Adams et al., 2017) and drive intense vertical circulations (J. R. Taylor et al., 2018). Mixed layer eddies are likely to be prevalent in regions where the mixed layer is deep and lateral gradients are sharp. They can arrest mixing and bring about spring mixed-layer stratification conditions earlier than with surface buoyancy forcing alone (Du Plessis et al., 2017). Further, the presence of submesoscale variability leads to the concentration of wind-driven near-inertial energy, enhancing the inertial wave shear-driven mixing below the base of the mixing layer (e.g., Klein et al., 2004; Meyer, Sloyan, et al., 2015; Jing et al., 2011). Large-scale inertial oscillations and submesoscale fronts may also induce transient modification of vertical stratification and thus turbulent mixing (L. N. Thomas et al., 2016). These observations point to the importance of the interplay of multi-scale physical processes in the Southern Ocean, a topic which is still largely unexplored.

There is increasing evidence that some submesoscale (∼ 1 km) processes in the surface mixed layer break the constraint of the large-scale quasi-geostrophic dynamics (i.e., the dominance of planetary rotation and vertical stratification), and trigger a variety of flow instabilities, such as inertial instabilities (Grisouard, 2018; Peng et al., 2020), symmetric instabilities (D’Asaro et al., 2011; L. N. Thomas et al., 2013), and ageostrophic baroclinic mixed layer instabilities (Boccaletti et al., 2007; Fox-Kemper & Ferrari, 2008). Unlike other surface processes that draw energy from atmospheric forcing, these instabilities extract either potential or kinetic energy from quasi-geostrophic flow (McWilliams, 2016), inject it into the smaller-scales of the fastest growing modes, induce secondary Kelvin-Helmholtz instabilities (J. R. Taylor & Ferrari, 2009) in the later stage, and finally mediate the transfer of energy from large-scale circulation to smaller scales through the forward cascade of energy (J. R. Taylor & Thompson, 2022). Several field studies with microstructure observations have confirmed the enhanced energy dissipation caused by this downscale transport of large-scale energy (D’Asaro et al., 2011; L. N. Thomas et al., 2016; Peng et al., 2020, 2021). Submesoscale frontal instabilities are especially relevant for the Southern Ocean because of the predominating atmospheric conditions of down-front winds and surface cooling.
Sea ice covers a large enough area of the Southern Ocean to have a large impact on air–sea interactions (§3.2.1). Fast ice provides a laterally rigid lid on the ocean that alters the mixing processes (Robertson et al., 1995; Stevens et al., 2009), from direct wind-forcing and surface wave breaking to ice-ocean frictional stresses associated with externally forced flows and tides (Albrecht et al., 2006). Sea ice strongly inhibits surface gravity waves and momentum fluxes from the wind thereby altering upper ocean mixing (§5.1.4 Ardhuin et al., 2020). However, the extent to which surface gravity waves and the associated dynamics, such as Langmuir circulations, are inhibited is dependent on the extent of the ice cover. Limited observations of air–sea–ice fluxes exist in the Southern Ocean. Observations from an air–sea flux mooring at the Polar Front (Ferreira Azevedo et al., 2022) found that 45% of surface energy penetrated the base of the mixed layer and suggest that even in the presence of sea ice, strong wind events may enhance mixing. Submesoscale activity and associated mixing can be enhanced under sea ice and in regions close to sea ice melt due to the existence of strong lateral density gradients. Observations at the edge of the Antarctic sea ice have revealed submesoscale eddies generated by the fresh water being stirred by the mesoscale eddies (see e.g., Giddy et al., 2021). Submesoscale activity has also been detected below the ice by observations from seal-based sensors (Biddle & Swart, 2020). Further, Gille et al. (2022) speculate that lateral density gradients resulting from heterogeneity in air–sea fluxes due to gaps between ice floes (Fons & Kurtz, 2019) could also lead to submesoscale-driven mixing.

Mixing at the face of, and underneath, ice shelves can be strongly influenced by tides (§5.2; Joughin & Padman, 2003; Padman et al., 2018). Tides generate increased turbulence in the layer of ocean adjacent to the ice, which modifies the temperature, salinity and density structure and leads to altered ocean circulation. This mixing is often modeled as a friction velocity, $u^* = C_D^{1/2} |u_s|$, where $C_D$ is a non-dimensional, quadratic drag coefficient that is $< 0.001$ where the ice base is smooth and up to approximately 0.2 under rough ice (Padman et al., 2018). The speed $|u_s|$ is the total flow at the ice–ocean interface, including both the general circulation and tides. The ice edge also induces substantial mixing, both in the wake but also in flow acceleration, depending on tidal conditions (Fer et al., 2012; Stevens et al., 2014). Within the cavity, it has been suggested that the interaction of tides and basal ice undulations might induce relatively high-frequency variability (Foster, 1983; Stevens et al., 2020), especially in the near-field of under-side basal crevasses (Lawrence et al., 2023). Under rapidly melting ice shelves, the freshwater outflow can generate currents that are much larger than the tidal currents. For example, the Pine Island Glacier has freshwater plume flow of up to 0.5 m s$^{-1}$ (Payne et al., 2007) and tidal currents of only a few centimetres per second (Robertson, 2013). Where cold ocean waters surround the ice shelf and melt rates are low, plume flows are much weaker than tidal flows. In these locations, the friction velocity and mixing will be dominated by the tidal currents.

Water mass transformation frameworks have revealed that wintertime mixing in the surface boundary layer of the Southern Ocean plays a key role in the diapycnal upwelling of Circumpolar Deep Water and the eventual formation of Antarctic Intermediate Water (Evans et al., 2018). Here, wintertime cooling and brine rejection during sea ice formation combine to weaken the stratification between the surface winter water and the Circumpolar Deep Water below. Mixing subsequently transforms the relatively warm and salty Circumpolar Deep Water into colder and fresher near-surface water. Through summertime warming and sea ice melt, this upwelled and transformed Circumpolar Deep Water eventually forms Antarctic Intermediate Water, likely through nonlinear thermodynamic processes (Evans et al., 2018).
4.3.2 Interior diapycnal mixing

Interior diapycnal mixing is widespread in the Southern Ocean due to the energetic internal wave environment (§5.3). The Southern Ocean is a region of strong wind-energy input into near-inertial motions (Alford, 2003). Both surface-generated near-inertial internal waves and bottom-generated internal tides and lee waves propagate into the interior, are shaped by interaction with other physical processes and subsequently break and generate diapycnal mixing. Interactions between the Southern Ocean’s energetic eddy field and the internal waves, in particular, lead to elevated diffusivity in the upper 2000 m of the ocean (Whalen et al., 2012, 2015, 2018). It is conceptually difficult to separate “interior” mixing from surface- and bottom-intensified mixing, both because of the surface/bottom boundary production of the waves that generate mixing and because there are many mixing hotspots associated with topography that extends high into the water column. Nevertheless, interior mixing below 2000 m depth, where interior diapycnal diffusivities are generally less than $10^{-4}$ m$^2$s$^{-1}$, drives interior watermass transformation and thus has important modulating impacts on the overturning circulation. These transformations and their importance to the Southern Ocean circulation are described in §2.5.

Internal waves can contribute significantly to small-scale ocean mixing, which itself contributes to large-scale processes such as watermass transformation and the overturning circulation (Munk & Wunsch, 1998). Wave breaking is the process through which internal waves dissipate. While they propagate, internal waves exchange energy with background mesoscale flows, such as currents, jets, and fronts, in what is called wave-mean interactions (Grimshaw, 1984), mesoscale eddies (wave–eddy interactions; Kunze, 1985; Cusack et al., 2020), or other internal waves (wave–wave interactions; McComas & Bretherton, 1977) resulting in the internal gravity wave continuum. The net energy flux can be from internal waves to their surroundings, while other times, from their surroundings to the internal waves (§5.3.2). Ultimately though, when internal waves reach high enough wavenumbers, they steepen and break through direct shear instability or convective overturning (Fig. 29), transferring their remaining energy into turbulence and diapycnal mixing (e.g., Eriksen, 1978; Fringer & Street, 2003; Nikurashin & Ferrari, 2010, where the latter is a Southern Ocean specific example). In the Southern Ocean, this wave breakdown is likely to happen during wave straining, capture and trapping in anticyclonic mesoscale eddies (Buhler & McIntyre, 2005) via the “inertial chimney” effect (Lee & Niiler, 1998; Z. Zhang et al., 2018), through wave scattering by the Antarctic Circumpolar Current and mesoscale eddies (Dunphy & Lamb, 2014), and in critical levels within jets (§5.3.2 Booker & Bretherton, 1967). This internal wave driven mixing can happen both locally, where internal waves originated, or remotely, when internal waves propagate far from their energy source. Such remote breaking and dissipation of internal waves is an important process for energy redistribution in the Southern Ocean, where the strong Antarctic Circumpolar Current has been documented to lead to the advection of internal waves through its fronts (Meyer, Polzin, et al., 2015), jets (Waterman et al., 2021), meanders and mesoscale eddies (Cyriac et al., 2023). Such modulation of the internal wave driven mixing landscape by the background mesoscale flow and associated wave-mean interactions might hold the answers to the discrepancy between the theoretical description of the Southern Ocean internal wavefield and the observed distribution of dissipation and mixing in recent studies (§7.2; Waterman et al., 2013; Sheen et al., 2013; Nikurashin et al., 2014; Cusack et al., 2017; Takahashi & Hibiya, 2019).

Globally, of the 2 TW of energy theorised to maintain the ocean stratification (Munk & Wunsch, 1998; de Lavergne et al., 2022), about 1.2 TW of energy is provided by internal waves generated from barotropic tides and geostrophic flows (Wunsch et al., 2004) with the remaining energy flux thought to come from the work done by wind on near-inertial motions (Alford et al., 2016). Uncertainty in these estimates is very large (§5.3), which leads to poor representation of wave-driven mixing in climate models (Jochum et al., 2013). Various estimates agree that much of the energy flux into lee waves occurs in the Southern
Ocean as expected given the uniquely deep-reaching nature of the Antarctic Circumpolar Current and relatively weak tidal flows (§5.2). Lee waves apply wave drag (§5.1.2) to the deep flows that generate them, which, in the Southern Ocean, are dominated by mesoscale eddies (Yang et al., 2018). The work done by the wave drag converts energy from the mesoscale eddy field into smaller-scale lee waves (Yang et al., 2018), which then transfer the energy further down to turbulence scales via direct wave breaking (e.g., Lefauve et al., 2015) or wave–wave interactions (e.g., Polzin, 2009).

Up to this point, §4.3.2 has focused on diapycnal mixing, which is the approximate vertical component of three-dimensional turbulence. Separating this mixing from the horizontal components set by (sub-)mesoscale stirring along isopycnals is convenient and is a common approach due to the different observations and methods used to estimate diapycnal and isopycnal diffusivities. However, it does not reflect the integrated three-dimensional nature of oceanic thermodynamical processes. A theoretical framework based on the temperature variance budget (Ferrari & Polzin, 2005; A. C. Naveira Garabato et al., 2016) establishes a balance between dissipation of variance by molecular mixing and the production of variance associated with mesoscale eddy-induced isopycnal stirring and with diapycnal mixing by small-scale turbulence acting on the large-scale mean state. The framework allows diapycnal and isopycnal diffusivities to be quantified from a small number of (temperature and velocity) microstructure measurements and provides new insight into the coupling between the zonal flow of the Antarctic Circumpolar Current and the meridional overturning circulation transport along sloping isopycnals. In Drake Passage (A. C. Naveira Garabato et al., 2016), the framework reveals that isopycnal stirring is strongly suppressed in the upper 1 km of Antarctic Circumpolar Current jets, consistent with earlier circumpolar work (A. C. Naveira Garabato et al., 2011). Intensified diapycnal mixing balances the meridional overturning in this upper 1 km, the lightest layer, and also in the densest layers of the Antarctic Circumpolar Current (A. C. Naveira Garabato et al., 2016). Both layers are close to the two primary sources of internal waves: wind-driven near-inertial oscillations and flow interactions with topography. Isopycnal stirring balances the overturning in the intermediate layers and upper Circumpolar Deep Water (A. C. Naveira Garabato et al., 2016). Application of the framework to only 10 microstructure profiles in the Brazil-Malvinas confluence (Oruër-Echevarría et al., 2021) reveals regional variations in the relative roles of diapycnal and isopycnal mixing. Observational campaigns, such as DIMES (Ledwell et al., 2011; Watson et al., 2013; Mackay et al., 2018), SOFINE (Waterman et al., 2013; Meyer, Sloyan, et al., 2015) and DEFLECT (Cyriac, Phillips, Bindoff, & Polzin, 2022) emphasize the importance of interactions between mesoscale variability, circulation and mixing for tracer transport (Mashayek, Ferrari, et al., 2017; Holmes et al., 2019). Greater use of microstructure observations will help unravel the roles of mesoscale, submesoscale and small-scale turbulent flows in governing the ocean’s circulation and water mass structure.

4.3.3 Bottom-intensified diapycnal mixing

Bottom-intensified mixing shapes Antarctic Bottom Water consumption, and underpins a key dependence of the abyssal circulation on both topographic roughness and large-scale topography (de Lavergne et al., 2017; Holmes et al., 2018; Polzin & McDougall, 2022). Bottom-intensified mixing is primarily generated by the breaking of lee waves and internal tides. Lee waves are generated through interactions between mesoscale flows and rough topography and are particularly prominent in the Southern Ocean due to the energetic mesoscale activity (Garabato et al., 2004; Nikurashin & Ferrari, 2010; Sheen et al., 2013; Meyer, Sloyan, et al., 2015; Gille et al., 2022). Internal tides also play an important role in this region (Johnston et al., 2015; Z. Zhao et al., 2018; Waterhouse et al., 2018; Vic et al., 2019).

Through nonlinear wave–wave interactions, these internal waves drive a down-scale cascade of turbulent energy leading to mixing (Nikurashin & Legg, 2011; Polzin, Naveira Garabato, Huussen, et al., 2014; Whalen et al., 2020) and a bottom-intensified profile of diffusivity
and buoyancy flux (Toole et al., 1994; Polzin et al., 1997; Waterhouse et al., 2014). As a consequence, a diapycnal transport dipole is established where there is downward transport (from light to dense water) in a stratified “bottom mixing layer” (often referred to as the stratified mixing layer) above the topography, and upward transport only within a much narrower “bottom boundary layer” where the turbulent buoyancy flux converges next to the topography (Fig. 16). The net diapycnal transport (or the consumption of Antarctic Bottom Water; Fig. 7) arises as a small residual of these larger upwelling and downwelling transformations (de Lavergne et al., 2016; McDougall & Ferrari, 2017; Ferrari et al., 2016; Polzin & McDougall, 2022). These mixing processes are shaped by both mesoscale and sub-mesoscale variability. Temporal variations in mixing associated with mesoscale eddy kinetic energy variations have been invoked to link bottom water overturning cell variability to wind forcing (Sheen et al., 2014; Broadbridge et al., 2016). Recent work also highlights the important role that near-bottom submesoscale processes play in maintaining the stratification, and, thus, the magnitude of the diapycnal transport dipole, in these bottom mixing layers (Ruan et al., 2017; Wenegrat et al., 2018; Callies, 2018; A. C. Naveira Garabato et al., 2019). These processes pose a particular challenge for modelling and observations given their small spatial scales and variability.
Figure 22. Schematic of gravity wave processes in the Southern Ocean including surface waves, internal waves and tides. At the surface, strong storm systems generate surface waves and (near-inertial) internal waves. The gravitational force of the moon and sun generate bulk motions of the water column (tides) that, in combination with other ocean flows, generate internal waves at the seafloor. The waves interact with other components of the Southern Ocean system. For example, surface waves are dissipated in the marginal ice zone, while internal waves may be trapped in eddies and currents, and/or drive diapycnal mixing in the ocean interior. Colour contours show a typical density field, ranging from lighter (dark orange) to denser (dark blue) waters.

5 Gravity waves

Gravity waves in the ocean are vertical perturbations of the fluid ocean against the restoring force of gravity, including displacements of the ocean surface (surface waves; §5.1), perturbations to the interior ocean stratification (internal waves; §5.3), and perturbations of the entire water column (tides; §5.2). These phenomena span from some of the smallest and fastest motions in the ocean in the case of surface waves (wavelengths of tens to hundreds of metres, periods of seconds), through intermediate length scales in the case of internal waves (horizontal wavelengths of kilometres to hundreds of kilometres), to motions that span ocean basins in the case of tides (thousands of kilometres). In all three cases, oceanic gravity waves are influenced by the Earth’s rotation — in addition to gravity — and are, therefore, more correctly termed “inertia-gravity waves”. These waves play a vital role in transporting energy and momentum throughout the ocean, thus supporting ocean mixing and circulation. Figure 22 provides a schematic overview of gravity waves in the Southern Ocean and their interactions with other components of the system. In this section, we present an overview of each class of gravity wave and its role in Southern Ocean dynamics.

For further details on gravity waves, readers may wish to peruse previous reviews in addition to the content herein. While there is no previous Southern Ocean specific review of surface waves, Young et al. (2020) collates over three decades of satellite altimeter and in situ buoy observations, to conduct a statistical study of seasonal variations, including extremes and spectral analysis, and highlights some of the unique aspects of Southern Ocean waves. Further, a series of articles (Squire et al., 1995; Squire, 2007, 2020) review the evolution in understanding of surface waves in the marginal ice zone (§3.2.2). For ocean tides, Pugh (2004) provides a detailed review of tidal dynamics and Stammer et al. (2014) reviews global tide models, with their Section 5.2 focusing on model performance in Antarctic seas.
In addition, Padman et al. (2018) describes ocean tide influences on the mass balances of the Antarctic and Greenland Ice Sheets. For internal waves, Polzin and Lvov (2011) provides a summary of the observed global ocean internal wave field and its explanation in terms of the nonlinear wave interactions (a subject not covered here). In addition, recent reviews have focused separately on internal waves generated at the ocean surface (L. N. Thomas & Zhai, 2022) and the seafloor (Musgrave et al., 2022), but with a global outlook.

5.1 Surface waves

Surface waves are generated by strong winds blowing in a coherent direction over the ocean surface, where the spatial extent of the region is known as the fetch (Donelan et al., 2006). Winds create random wave fields that consist of a spectrum of frequencies and directions (Holthuijsen, 2010). Surface waves are dispersive, such that lower frequency components (i.e., longer waves) travel faster than higher frequency components. They also have the property that the speed of energy transport of a frequency component (the group speed) travels slower than corresponding wave crests (the phase speed). In the open ocean (i.e., where the water is deep compared to the wavelength), the group speed is half the phase speed (Holthuijsen, 2010). In the Southern Ocean, surface gravity waves have typical periods of 10–12 s and wavelengths of 150–230 m.

Surface waves are directly generated at wavelengths such that their phase speeds are resonant with the prevailing surface winds. Only waves with phase speeds similar to or slower than the wind speed are directly generated (Young, 1999). Nonlinear wave–wave interactions redistribute part of the spectrum towards low frequencies, downshifting the spectral peak. Thus, the dominant wave form stretches and, because of dispersion (Young, 1999), accelerates as its height grows (Tulin & Waseda, 1999; Young, 1999; Toffoli et al., 2017). A fraction of energy also cascades towards higher frequencies, forming a decaying high-frequency spectral tail (Zakharov et al., 2012; Onorato et al., 2002; Toffoli, Onorato, et al., 2010; Toffoli et al., 2017; Romero et al., 2012; Young et al., 2020).

After long fetches, waves reach full development, becoming independent from local winds. Further development of the wave field is associated with nonlinear interactions (Young, 1999). As a consequence, the “wind sea” (i.e., a wave field being acted on by winds) evolves into more regular wave fields that radiate along multiple directions from the generation area. These so-called “swells” disperse across the Indian, Pacific, and South Atlantic Oceans (Semedo et al., 2011).

5.1.1 Southern Ocean surface waves

The Southern Ocean possesses a unique surface wave climate due to the absence of large land masses, which allows circumpolar-scale fetches and persistently strong westerly (i.e., blowing from west to east) winds, including the notorious ‘roaring forties’, ‘furious fifties and ‘screaming sixties’ (Lundy, 2010). Westerlies give rise to some of the fiercest surface waves on the planet (Barbariol et al., 2019; Vichi et al., 2019; Young et al., 2020; Derkani et al., 2021; Alberello et al., 2022). The wave height climate mirrors the distribution of wind speeds, with a uniform distribution of waves across the region and the seasons. The “significant wave height”, i.e., the average height of the highest one-third of the waves experienced over time (Young, 1999), is in excess of 3.5 m in summer (Fig. 23a) and 5 m in winter (Fig. 23b), according to model hindcasts and satellite observations (Young et al., 2020; Schmale et al., 2019; Derkani et al., 2021). Long term in-situ observations at several locations reveal that extreme events, with significant wave heights greater than 10 m, can occur over winter with a probability of approximately one in one-thousand (Rapizo et al., 2015; Young et al., 2020).

At short fetches, the wave spectrum is narrow banded (Young et al., 2020) and the wave form is steep, facilitating occurrence of highly nonlinear dynamics (Janssen, 2003; Onorato
et al., 2009). Laboratory experiments in a circular wave flume that mimic the unlimited fetch conditions in the Southern Ocean suggests that nonlinear dynamics have the potential to fully develop, causing individual waves to destabilize and grow significantly taller than the background sea state. In exceptional circumstances, this leads to so-called “rogue waves”, which have heights greater than two times the significant wave height (Toffoli et al., 2017). Exceptional maximum individual wave heights exceeding 19 m have been reported (Barbariol et al., 2019), although these are not necessarily rogue waves.

Observations around the Southern Ocean indicate that very broad directional distributions are common in the region, with energy spreading across a range up to ±80° around the mean wave direction (Young et al., 2020; Derkani et al., 2021). On occasions, this is the signature of chaotic sea states, where multiple (independent) wave systems, such as wind seas plus swells coexist (Aouf et al., 2020; Khan et al., 2021; Derkani et al., 2021; Alberello et al., 2022). Theory, numerical simulations and experiments have demonstrated that these multi-system seas accelerate development of nonlinear dynamics, further contributing to the occurrence of large amplitude waves (Onorato et al., 2006; Toffoli, Bitner-Gregersen, et al., 2011).

5.1.2 Wave breaking

Waves grow under the forcing of wind and highly nonlinear instabilities until they ultimately break in the form of whitecaps (Babanin et al., 2007; Toffoli, Babanin, et al., 2010; Toffoli et al., 2017), when the ratio of wave height to wavelength is \( \approx 0.14 \) (Fig. 24; Toffoli, Babanin, et al., 2010). Theoretically, this ratio occurs when water particle velocities at the wave crest exceed the phase velocity, a condition which corresponds to a wave profile having an angle between the upwind and downwind sides of the wave crest of 120° (Stokes, 1880; Michell, 1893). Whitecaps can be explained as pressure pulses on the sea surface just downwind of the wave crest that act against wave growth (Hasselmann, 1974), dissipating excessive wind input and, subsequently, transferring it to the subsurface in the form of turbulent mixing (§4.3.1; Terray et al., 1996). However, breaking-induced turbulence decays while maintaining energy levels.
Figure 24. Schematic of a breaking surface gravity wave. The wave propagates in the direction of the wind and grows with time until it becomes too steep and breaks. Wave breaking induces near surface turbulence, which generates air bubbles and entrains them into the sub-surface ocean, mediating air–sea fluxes of momentum, energy, moisture and biological constituents with the ambient atmosphere. Turbulent oscillatory motion (from both breaking and non-breaking waves) drives vertical mixing (blue spirals) through the water column to a depth comparable to the wavelength, contributing to the mixed ocean surface layer.

...rapidly in depth with distance from the surface and the contribution to ocean mixing is confined to a sublayer with depths comparable with the wave height (Rapp & Melville, 1990). Nevertheless, there is theoretical and experimental evidence that the wave oscillatory flow can become turbulent even in the absence of breaking (Babanin, 2006; Alberello, Onorato, Frascoli, & Toffoli, 2019). Hence, waves are capable of directly contributing to mixing throughout the water column, until depths comparable to half of the wavelength (i.e., down to about 100 m; Toffoli et al., 2012).

Besides dissipation, whitecaps drive air–sea interaction processes through airborne droplets (Monahan et al., 1986; Landwehr et al., 2021). Generated and entrained sub-surface by whitecaps, bubbles rise to the surface and burst, forming film droplets or jets of daughter droplets (Fig. 24). If the wind shear is sufficiently intense, larger droplets known as “sea spray” are torn off the surface of (breaking) waves (Veron, 2015). Once ejected, spray drops are transported and dispersed in the marine atmospheric boundary layer, in which they interact and exchange momentum, heat, moisture and biological and chemical constituents with the ambient atmosphere (Humphries et al., 2016; Schmale et al., 2019; I. Thurnherr et al., 2020; Landwehr et al., 2021). There is evidence that marine aerosols generated from whitecaps are an important source of cloud condensation nuclei and cloud formation in the Southern Ocean (Schmale et al., 2019; Landwehr et al., 2021). Large sea spray particles do not dissolve entirely while in the atmosphere, but they return to the ocean with lost or gained momentum, closing the loop of air–sea interaction (Veron, 2015; Landwehr et al., 2021).

5.1.3 Influence of mesoscale currents

Mesoscale ocean currents (approximately ten to one hundred of kilometres) influence generation and propagation of surface waves. Wind forcing depends on adjustments of the atmosphere relative to current speed (Ardhuin et al., 2017), which accelerates wave growth when wind opposes the current or decelerates it when wind follows the current. Theoretical
and experimental evidence (Onorato et al., 2011; Toffoli, Cavaleri, et al., 2011; Toffoli, Waseda, et al., 2015) shows that notable changes to wave height, frequency and direction can occur when the current speed relative to the phase speed is greater than 0.02. Changes in wave frequency are the result of a Doppler shift, which compresses wavelengths if waves oppose the current or stretches it if waves follow the current. Variation of wave direction (i.e., refraction) relates to changes in phase speed and results in the broadening of the directional distribution (Toffoli, Cavaleri, et al., 2011; Toffoli, Waseda, et al., 2015; Rapizo et al., 2016). The wave height increases in opposing currents due to energy bunching (a phenomenon similar to wave shoaling in coastal waters; Holthuijsen, 2010), directional focusing due to refraction and nonlinear wave interactions, which can then lead to dissipation processes (Onorato et al., 2011; Toffoli, Cavaleri, et al., 2011; Toffoli, Waseda, et al., 2015; Rapizo et al., 2015). In contrast, wave height decreases in following currents due to relaxation induced by reduction in the wave energy transport velocity (i.e., the group velocity).

Velocities in the jets of the Antarctic Circumpolar Current (§§ 2.1, 4.1) exceed 0.75 m s\(^{-1}\) (Derkani et al., 2021), but its main effect on the wave field is only refraction, as it flows predominantly in the direction of the waves. Therefore, the Antarctic Circumpolar Current helps maintain the broad directional distribution observed in the region (Derkani et al., 2021; Young et al., 2020). As waves propagate along the current, the wave height is attenuated, although this effect is small (Derkani, 2021; Rapizo et al., 2015). More substantial interactions are reported at the upper boundary of the Indian Ocean sector, where large swells from Antarctica interact with the more intense Agulhas current, forming large amplitude waves and, often, rogue waves (White & Fornberg, 1998).

5.1.4 Attenuation, dissipation and scattering by sea ice

![](image)

**Figure 25.** Observations (symbols) and power-law best fits (with exponent \(n\); lines) of the wave attenuation coefficient in the marginal ice zone from the Arctic (gray) and Southern Ocean (colours). Adapted from Meylan et al. (2018) to include results by Montiel et al. (2022) of low ice concentration/shorter waves (blue), mid-range concentration and wave period values (red), and high concentration/longer waves (green).
As surface waves penetrate into the marginal ice zone (§3.2.2) from the sea ice free open ocean, the sea ice loads acting on the ocean surface change the wave propagation characteristics. The ice-loaded ocean surface medium is characterised by the local sea ice conditions, i.e., concentration, thickness, ice type, floe size distribution, etc., which are in turn related to the wave conditions (§3.2.2). A collection of in-situ and remote sensing observations (originally from the Arctic but, more recently, also from the Southern Ocean) provide evidence that ocean wave energy decays exponentially with distance travelled though the MIZ and that the rate of attenuation increases with wave frequency (Squire & Moore, 1980; Wadhams et al., 1988; Kohout et al., 2014; Meylan et al., 2014; Stopa et al., 2018; Montiel et al., 2018; Kohout et al., 2020; Montiel et al., 2022; Alberello et al., 2022). The observations suggest that the rate of exponential attenuation, which is known as the attenuation coefficient, has a power-law relationship with wave frequency (Fig. 25; Meylan et al., 2018). Understanding how the attenuation coefficient emerges from the underlying dynamic processes has been the main focus of ocean waves–sea ice interactions research over the past half century (Squire et al., 1995; Squire, 2007, 2020; Golden et al., 2020).

In situations where the floe sizes are comparable to the wavelengths, the ice floes scatter the waves over the directional spectrum (Fig. 8). Scattering is an energy-conserving process but an accumulation of scattering events causes waves to attenuate over distance (Squire, 2007, 2020). Much theoretical work has attempted to describe wave attenuation due to linear wave scattering in the MIZ, using phase-resolving multiple-scattering theory in one horizontal dimension (Kohout & Meylan, 2008; Bennetts & Squire, 2012b) or two dimensions (Bennetts & Squire, 2009; Peter & Meylan, 2010; Bennetts et al., 2010; Montiel et al., 2016). There have also been theories proposed to include attenuation due to scattering in phase-averaged wave transport models, using energy sink terms (Dumont et al., 2011; T. D. Williams et al., 2013a, 2013b; Mosig et al., 2019), a Boltzmann-interaction term (Meylan et al., 1997; Meylan & Masson, 2006; Meylan & Bennetts, 2018; Meylan et al., 2020) or a diffusion term (X. Zhao & Shen, 2016). Wave scattering through random fields of ice floes results in (i) exponential attenuation at a rate that increases with frequency, qualitatively consistent with observations, and (ii) broadening of the directional spread, so that deep into the MIZ the directional wave spectrum becomes isotropic (Wadhams et al., 1986; Meylan et al., 1997; Bennetts et al., 2010; Montiel et al., 2016; Squire & Montiel, 2016). Scattering models show reasonable agreement with historical measurements from the Arctic in the mid-frequency regime where linear scattering theory is valid (Kohout & Meylan, 2008; Bennetts et al., 2010; Bennetts & Squire, 2012a; Squire & Montiel, 2016).

Measurements of surface waves in the Antarctic MIZ have been made over the past decade, predominantly using specially designed wave buoys deployed on the surface of ice floes (Kohout et al., 2014; Meylan et al., 2014; Kohout et al., 2020; Montiel et al., 2022), and recently by a stereo-camera system on an icebreaker (Alberello et al., 2022). The floe sizes during the observations were typically much smaller than wavelengths, e.g., pancake ice (§3; Alberello, Onorato, Bennetts, et al., 2019), for which dissipative processes are likely to be the main contributors to wave attenuation. Dissipative processes can broadly be separated into turbulent ocean processes and viscous ice processes. Turbulence through wave–ice interactions occurs as a result of the differential velocity between the solid ice boundary and the water particle orbital velocity (Voermans et al., 2019). A turbulent boundary layer is generated at the basal surface of ice floes, which can be enhanced by the ice surface roughness (skin friction) and the presence of vertical ice features, e.g., ice floe edges or pressure ridges, further enhancing flow separation (form drag; Kohout et al., 2011). Turbulence also occurs in overwash on upper surfaces of floes, resulting in wave energy dissipation (Bennetts et al., 2015; Bennetts & Williams, 2015; Toffoli, Bennetts, et al., 2015; Nelli et al., 2017, 2020). Sea ice covers have been modelled as viscoelastic materials, such that they experience viscous dissipation when strained by ocean waves (Keller, 1998; R. Wang & Shen, 2010; Mosig et al., 2015). For instance, unconsolidated grease or brash ice (§3) dissipates wave energy through non-recoverable, shear stress-induced viscous deformations (Weber, 1987; G. Sutherland et al., 2019). Quantifying these dissipative processes is challenging, as they
depend on temperature, brine volume fraction and ultimately the micro-structure of the ice cover (Timco & Weeks, 2010). In a more heterogeneous ice cover (e.g., pancake ice) wave energy dissipation is more likely to be governed by eddy-generating floe–floe collisions (Shen & Squire, 1998; Bennetts & Williams, 2015; Yiew et al., 2017; Rabault et al., 2019; Herman et al., 2019).

5.2 Tides

Tides are a ubiquitous feature of the global ocean. Gravitational dynamics of the Earth–Moon–Sun system, combined with the Earth’s rotation, cause oscillations of ocean height and currents at precise periods, dominated by diurnal (daily) and semidiurnal (twice daily) tidal constituents. Here we use the term “tide” to exclusively describe the barotropic or surface tide, as opposed to “internal tides”, i.e., tidally generated internal waves, which are described in §5.3.

Tides are both noise and signal. Tide heights must be removed when processing satellite altimetry data over the ocean, sea ice and ice shelves to reveal trends in sea surface height and associated currents (e.g., Armitage et al., 2018), sea ice freeboard, and ice shelf thickness and mass change (e.g., Adusumilli et al., 2020). Tidal signals must also be removed to reveal background synoptic and mean flows in observations that are dominated by tides in many regions. While this procedure is straightforward for tide-resolving data from mooring time series, it is more complex for intermittent current measurements from ship-based and glider data. Here we focus on tides as signal, since they provide a substantial fraction of the total kinetic energy in the Southern Ocean, with known effects at all scales from turbulence (§4) to large-scale circulation (§2).

5.2.1 Southern Ocean tides

Throughout most of the global ocean, tides exist as propagating barotropic waves. These waves have spatial scales comparable to ocean basins and their amplitude depends on the global distribution of continents and bathymetry. These propagating waves are relatively straightforward to constrain in inverse models using in situ and satellite data (e.g., Egbert & Erofeeva, 2002; Lyard et al., 2006). However, the largest tidal currents around Antarctica are associated with diurnal-band, topographically trapped vorticity waves along the shelf break. These waves are a specific, tidally-forced case of the coastal trapped waves (§5.3.3). Observations and models of diurnal topographic vorticity waves (e.g., J. H. Middleton et al., 1987; Semper & Darelius, 2017; Skardhamar et al., 2015) show that they can have short spatial scales, are poorly constrained by sea surface height data, are extremely sensitive to topographic variability, stratification and mean flows, and produce strongly depth-varying currents. Predicting these currents is a difficult modelling problem, especially at the typical coarse grid scales of global climate models. When present, these waves have a profound effect on cross-slope transport of ocean heat, mean flows through tidal rectification (Makinson & Nicholls, 1999; Flexas et al., 2015), and the volume flux and hydrographic characteristics of Dense Shelf Water and Antarctic Bottom Water outflows (e.g., Padman et al., 2009).

The ratio of tidal frequency to inertial frequency has a strong influence on tidal dynamics. For diurnal tides at high latitudes, the tidal frequency is about half the inertial frequency, leading to the potential for generation of strong, topographically trapped vorticity waves (e.g., J. H. Middleton et al., 1987) that are largest in regions of rapid changes in water depth including the shelf slope and ice-shelf fronts. The strongest semidiurnal constituent, M2, has a critical latitude (where the inertial frequency equals the tidal frequency) of about 74.5°. M2-baroclinic tides (§5.3) cannot freely propagate south of this latitude, and friction at the bottom boundary and any ocean–ice boundary can lead to large changes in boundary layer thickness in these regions.
Figure 26. Schematic of the primary roles of tides in the Southern Ocean system. Under the ice shelf, tidal currents generate friction that modifies hydrographic properties of the water column, influencing the basal melting of the ice shelf. At the ice front and shelf break, rectified tidal currents modify water mass transport along and across these topographic barriers. Over the continental shelf and slope, tidal currents modify sea ice production and concentration. Stress at the base of landfast sea ice affects melting, and mixing controls on surface mixed layer depth. Mixing and rectification of tidal flows alters the production of Antarctic Bottom Water (AABW). Farther north, tidal flows over steep and rough topography of mid-ocean ridges generates internal (baroclinic) tides that can drive mixing in the ocean interior. Baroclinic tides may also be generated over the continental slope.

The change in sea surface height due to tides is not dynamically important for most of the Southern Ocean, although it does play a role in ice sheet weakening through crevasses formation by flexure at grounding lines of ice shelves and tidewater glaciers (Padman et al., 2018). Instead, the principal role of the tide is through the interactions of tidal currents with other components of the system, including as a source of mixing (§ 4.3), divergent stresses on sea ice (Padman & Kottmeier, 2000; Heil et al., 2008), and basal melting of ice shelves (Fig. 26; Richter et al., 2022). Tides link processes ranging from the smallest time and space scales of mixing to the global scales of continents, ocean basins and ice sheets that set the spatial distribution of tidal currents (e.g., Figs. 1b and 9b of Padman et al., 2018). The largest tidal currents around Antarctica are found along the shelf breaks of the Ross and Weddell seas, and under Ronne Ice Shelf. Along the Northwest Ross Sea shelf break, maximum spring tidal currents can exceed 1 m s^{-1} (Whitworth & Orsi, 2006). Tides in the Pacific sector are dominated by diurnal variability, while semidiurnal tides dominate elsewhere (e.g., Fig. 1c of Padman et al., 2018).

5.2.2 Tide measurements and models

There are relatively few in situ tide height measurements in the Southern Ocean (M. A. King & Padman, 2005). High quality tidal records are limited to a few long-duration coastal tide gauges and bottom pressure recorders. While many shorter records exist, these are poorly suited to tidal prediction. However, recent deployments of Global Navigation Satellite System (GNSS) receivers on ice shelves have provided high quality tide records greater than one year long (e.g., Ray et al., 2021). Additional data come from satellite altimetry (reviewed in Section 2.2.2 of Padman et al., 2018), both for large ice shelves and the open ocean (Stammer et al., 2014). Tidal currents are measured with current meters that are typically deployed for one or more years. However, individual moorings rarely cover
the full depth of the water column meaning that depth-averaged currents are difficult to evaluate.

Given the paucity of high quality in situ data, our modern knowledge of Southern Ocean tides comes primarily from ocean tide models. Barotropic models solve the depth-integrated equations of motion and provide depth-averaged (“barotropic”) currents. For example, the global solutions reviewed by Stammer et al. (2014) or regional models such as CATS2008 (S. L. Howard et al., 2019). These models may be based entirely on dynamics (with open boundary conditions applied for regional models) or inverse models constrained by assimilation of ocean height data including in situ measurements and satellite altimetry. The accuracy of barotropic models in the Southern Ocean, especially in the Antarctic coastal seas, is typically lower than at lower latitudes because the best satellites for tidal studies (TOPEX/Poseidon and Jason) only sample to about 66°S and, for these and other satellites with higher-latitude orbits, the presence of sea ice and ice shelves complicates retrieval of the ocean tidal signal.

5.2.3 Tidal rectification

Nonlinear interactions between tidal flows and a sloping seafloor (such as the continental shelf), in the presence of planetary rotation and spatial variations in tidal amplitude, can lead to the generation of a time-averaged mean flow, in a process known as “tidal rectification” (Loder, 1980; I. Robinson, 1981). These time-averaged flows have speeds of approximately 10–15% of the tidal current. Observations and models suggest that rectified tidal flows across the Northwest Ross Sea outer continental shelf play an important role in the Antarctic Bottom Water export from this region (Fig. 10 of Padman et al., 2009). Makinson and Nicholls (1999) implicated tidal rectification as playing a key role in the ventilation of the ocean cavity under Filchner-Ronne Ice Shelf. Numerical modelling studies (Flexas et al., 2015) have also shown that tidal rectification-induced volume flux convergence is essential to simulate a realistic Antarctic Slope Front and Current (§2.2).

In locations where tidal currents are comparable to mean flows, they can also modify those mean flows through changing the time-averaged stress at the seafloor. This tidal rectification is distinct to that discussed above since it involves the modification of existing mean flows, rather than the interaction of the tide with topographic gradients to generate new mean flows (Loder, 1980). Robertson et al. (1985) postulated that strong tides around the perimeter of the Weddell Sea could significantly reduce the transport of the Weddell Gyre through tide-induced weakening of mean flows via this rectification mechanism. A similar response is expected in other locations where benthic tidal currents are significant, notably in the southern limb of the Ross Gyre.

5.2.4 Internal tide drag

Internal tide drag is the periodic force exerted on the surface tide when it interacts with seafloor topography to generate internal waves. This effect is the dominant tide–topography interaction in the deep, open ocean where tidal flows are weak (a few centimetres per second) and turbulent drag, which dominates in regions of strong tidal flow on continental shelves, is negligible. The key role of internal tide drag was directly identified with the advent of satellite observations and associated inverse models indicating that approximately 30% of energy loss from the surface tide occurred in the open ocean (Egbert & Ray, 2000). Internal tide drag was subsequently implemented in forward-running tide and ocean models (Jayne & St. Laurent, 2001). It is now recognised that internal tide drag is crucial in setting the amplitude of the surface tide (Buijsman et al., 2015; Arbic et al., 2018), and, therefore, also feeds back on the strength of internal tide generation (Ansong et al., 2015) (see (§5.3.1). Recent work has shown that this internal tide drag is not purely a drag force, but also exhibits an out-of-phase force component, analogous to the spring in a harmonic oscillator, which can both damp and, in certain resonant configurations, amplify the sur-
face tide (Shakespeare et al., 2020). This out-of-phase force component dominates when sub-inertial topography-trapped internal tides are generated (i.e., poleward of the critical latitude; §5.3.3) and, therefore, may be particularly important in the Southern Ocean.

5.3 Internal gravity waves

Internal waves are a key component of the ocean system. They transfer energy from large scale motions to small scale turbulence, making them a major source of interior ocean mixing. The mixing generated by internal waves is one of the drivers of large scale ocean circulation and plays an important role in biological and physical interactions, including the transport of nutrients and larvae. Internal waves also transport momentum into the ocean from the boundaries, thereby directly forcing the eddying and larger-scale circulation. They are generated when the ocean density field is perturbed and can be identified as oscillations of these different layers of the stratified ocean interior. Internal waves have vertical length scales from a few meters to 2 km, horizontal length scales from a few meters to hundreds of kilometers, horizontal group velocities of 10–100 mm s$^{-1}$, amplitudes from meters to hundreds of meters, and periods from several minutes to a day (Thorpe, 2007; Kantha & Clayson, 2000).

Internal waves originate primarily at the ocean’s upper and lower boundaries. They are forced by wind stress fluctuations at the surface, and at the seabed by tides and mesoscale flows interacting with rough topography. Observations of near-inertial wave energy propagation from the mixed layer into the ocean interior suggest that wind-generated internal waves are an important part of the ocean mixing budget in the Southern Ocean (Waterman et al., 2013). The Southern Ocean has a deep-reaching mesoscale flow, sometimes referred to as the “mean flow” or “background flow”, which is a mix of strong currents such as the Antarctic Circumpolar Current and associated jets, meanders and mesoscale eddies (§§2.1,2.5). The interaction of this deep-reaching mesoscale flow with the seafloor is a major source of topographic internal waves in the Southern Ocean (Nikurashin & Ferrari, 2011, 2013). New maps of internal tide-induced sea surface height perturbations derived from repeat-orbit satellite altimetry (Zaron, 2019) have revealed energetic internal tides near the Kerguelen Plateau, Macquarie Ridge and Drake Passage, consistent with previous modelling studies (Simmons et al., 2004). Vertical displacement variance at 1000 m depth measured with Argo profilers, has uncovered similar hotspot regions, particularly in the Kerguelen Plateau and Drake Passage regions (Hennon et al., 2014).

5.3.1 Internal wave generation in the Southern Ocean

The Southern Ocean storm track, sometimes called the ‘roaring 40s’, centred on 40°S, is associated with high wind work and is a key source of near inertial waves (Simmons & Alford, 2012). Wind blowing at the local inertial frequency band can force inertial motions through resonance in the ocean surface mixed layer (D’Asaro, 1985; Alford et al., 2016; L. N. Thomas & Zhai, 2022). Those inertial motions lead to convergence and divergence at the stratified base of the mixed layer. This pumping generates internal waves close to the local inertial frequency (or Coriolis frequency) everywhere in the ocean except at the equator, which we call “near-inertial waves” (Fig. 27). The resonant frequency, or effective local Coriolis frequency at which near-inertial waves are generated is modified by the relative vorticity of the background flow (e.g., Kunze, 1985; Schlosser, Jones, Bluteau, et al., 2019, for the Southern Ocean). Near-inertial waves are mostly generated from wind (other mechanisms are discussed below), and propagate almost exclusively equatorward, since the inertial frequency decreases with latitude (Garrett, 2001; Chiswell, 2003; Alford & Zhao, 2007). They are blocked from propagating poleward, except in strongly sheared currents (Jeon et al., 2019), since their frequency would become sub-inertial, typically within a single wavelength. Globally, most of the ocean’s kinetic energy (Leaman & Sanford, 1975; Garrett, 2001; Wunsch et al., 2004) and vertical shear (Alford et al., 2016) is in the near-inertial band, standing apart from the rest of the internal wave spectrum (L. N. Thomas
Figure 27. Schematic of near-inertial waves generation, propagation and dissipation. Storms generate inertial oscillations in the ocean mixed layer which drive horizontal convergences and divergences that lead to vertical velocities. These pump the base of the mixed layer generating internal waves near the local inertial frequency (1–1.2 $f$) that have counterclockwise polarization in the Southern Ocean. High mode near-inertial waves propagate downward and equatorward and tend to break locally due to high shear. Low mode internal waves propagate further equatorward. The interactions between near-inertial internal waves with other internal waves and with the background mesoscale flow are not represented here. Figure adapted from Alford et al. (2016).
The second major source of internal waves is via the interaction of ocean flows with the rough seafloor (Musgrave et al., 2022) and the Southern Ocean is a hotspot for a certain type of these topographically generated internal waves known as ‘lee waves’. When a fluid parcel is lifted up and over a topographic obstacle at sufficient speed, the restoring buoyancy force from the stratification initiates an oscillation (internal wave) which radiates energy away from the seafloor (Fig. 28). The ocean flow doing the lifting is a combination of eddies, jets and other currents (§ 4.1), which are essentially steady on the timescale of waves (< 1 day) and the barotropic tide (§ 5.2), which varies on sub-daily timescales (frequency $\omega$). Assuming a background mesoscale flow speed of $U$ and topographic wavenumber $k$, generation of freely-propagating topographic internal waves can only occur in the regime where the intrinsic frequency is between the inertial frequency, $f$, and buoyancy frequency $N$; i.e., $|f| < |\omega + kU| < |N|$. Therefore, barotropic tides (through frequency $\omega$) and mesoscale flow (through speed $U$) conspire in the generation of internal waves at topography (Bell, 1975; Shakespeare, 2020). The two end members of topographic internal waves are steady lee waves (when there is no tidal flow) and pure internal tides (when there is no quasi-steady flow). Steady lee waves are only generated at very small scale topography ($|f| < k < |N/U|$), which for typical deep Southern Ocean conditions ($U = 0.1–0.2 \text{ m s}^{-1}$, $f = 1 \times 10^{-4} \text{ rad s}^{-1}$, $N = 1 \times 10^{-3} \text{ rad s}^{-1}$) restricts $2\pi/k$ to topographic scales of 0.5–10 km. Consequently, the presence of small-scale topography critically determines the geographical location of lee wave generation. By contrast, pure internal tides are only generated where $\omega > f$ (equatorward of ~ 74.5° for semi-diurnal, and ~ 28° for diurnal) at large scale (small $k$) topography where the influence of the background mesoscale flow is negligible. In intermediate regimes, topographic internal waves exist as “Doppler shifted internal tides” (Shakespeare, 2020) but most studies have focused only on the two limiting cases.

The energy flux into topographic internal waves is determined primarily by the stratification, $N$, at the ocean bottom, topographic spectrum and flow speeds: $E \sim \rho_0 kN^2U^2h^2$ where $k$ is the mean topographic wavenumber, $h$ the root-mean-squared height, and $U$ the appropriate tidal or quasi-steady flow speed (e.g., Garrett & Kunze, 2007). This scaling only applies in the so-called “intermediate frequency limit”, where $|f| \ll |\omega + kU| \ll |N|$. Thus, the weak stratification typical of the Southern Ocean at the depth of prominent bathymetric features tends to limit the production of internal waves, but this is somewhat counteracted by the presence of unusually rough and large amplitude topography, and deep-reaching, intense eddying flows. However, it is not a simple matter of additional lee wave generation in the Southern Ocean compensating for reduced internal tide generation, since the fates of these waves are likely to be very different. While lee waves are confined within their generating flow (e.g., jet, eddy, meander), internal tides can freely propagate into different flow regimes (Shakespeare, 2020). We first consider the magnitude of pure internal tide generation at large scales, before discussing the small-scale limit where both internal tides and lee waves are generated.

For the dominant M2-tidal constituent, total low-mode internal tide generation in the Southern Ocean has been estimated from baroclinic tide models to be 0.15 TW (compared with 0.87 TW globally) with almost all the energy flux occurring at three locations: Macquarie Ridge, Kerguelen Plateau, and in the vicinity of Drake Passage (Table 3 of Simmons et al., 2004). However, other modelling (Padman et al., 2006) suggests these values may be significantly overestimated, and that the calculation may be highly resolution dependent.

At horizontal scales of ~ 10 km or less, and especially in the Southern Ocean, the seafloor is dominated by features known as “abyssal hills” (Goff, 1991), which are not resolved in the bathymetric datasets or large scale models. However, spectral representation of this topography (Goff, 2010; Goff & Arbic, 2010), together with numerical model estimates...
Figure 28. Schematic of internal lee wave generation and the associated vertical transfer of horizontal momentum flux via form stress across isopycnal layers. The pressure is increased on the upstream side of the hill ($+\Delta P$) and decreased on the downstream side ($-\Delta P$), resulting in a force from the fluid on the hill. The breaking of the wave at a critical level drives turbulent mixing and deposition of the wave momentum, resulting in a net force on the background mesoscale flow. For lee waves, this force always acts to decelerate the flow.

of $N$ and eddying flow $U$, may be used to estimate internal wave generation at abyssal hills. Globally, an additional $M_2$ internal tide energy flux of 0.03–0.1 TW is thought to generated, but only perhaps 10% of this flux occurs in the Southern Ocean (Melet et al., 2013; Shakespeare, 2020). Many authors (A. C. Naveira Garabato et al., 2013; Scott et al., 2011; Nikurashin & Ferrari, 2010; Wright et al., 2014; Yang et al., 2018; Shakespeare, 2020) have also used linear theory (following Bell, 1975, but often with modifications to account for nonlinear effects) to calculate rates of lee wave generation globally. Predictions vary from 0.05 to 0.85 TW, with the majority of this energy flux usually concentrated in the Southern Ocean. The huge range of estimated energy flux for small-scale internal tide and lee wave generation is due to the extreme degree of uncertainty in numerical model estimates of both bottom stratification and eddying flow speeds at the seafloor, as well as a paucity of observations to constrain the models.

Other sources of internal waves in the Southern Ocean are the relative motion of sea ice across the upper ocean through the shape of the under-ice surface (McPhee & Kantha, 1989), ice floe motions (Waters & Bruno, 1995), and ice tongues and ice shelf basal variability. Internal wave generation under sea ice is controlled by sea ice roughness, sea ice concentration and wind forcing (S. T. Cole et al., 2018). While such sea ice generated internal waves have been reported in the Arctic (S. T. Cole et al., 2014), there are currently few direct observations of internal waves under Antarctic sea ice and ice shelves, which are limited to internal tides for example in mooring data (e.g., S. Howard et al., 2004). The magnitude of energy fluxes from these generation mechanisms, which are harder to observe and model, and their relative prevalence are unknown. Additional internal waves generation mechanisms that are not specific to the Southern Ocean are adjustment processes (e.g., geostrophic adjustment) at fronts and eddies (Gill, 1984; Alford et al., 2013; Nagai & Hibiya, 2015; Rijnsburger et al., 2021) and frontogenesis (Shakespeare, 2019).
5.3.2 Influence of geostrophic turbulence on internal waves

The interaction of the strong Southern Ocean mesoscale flow with the seafloor gives rise to the emission of internal waves that possess a net momentum directed mostly against the flow (Bell, 1975; Nikurashin & Ferrari, 2011; A. C. Naveira Garabato et al., 2013; Shakespeare & Hogg, 2019; Shakespeare, 2020). This momentum is transported by the waves and deposited where they break and dissipate, leading to a net force on the fluid (Eliassen, 1961; Bretherton, 1969; Andrews & McIntyre, 1978). In the case of lee waves, this force is often termed the “lee wave drag”, which plays a significant role in Southern Ocean dynamics (Fig. 28; A. C. Naveira Garabato et al., 2013). The wave dissipation may be triggered by various mechanisms including shear instabilities (Fig. 29), wave saturation, wave–wave and wave–mean interactions.

Wave–mean interactions encompass all mechanisms of interactions that are the result of wave propagation through gradients in velocity and density induced by eddies, jets or any other currents. For example, lee waves propagating upward and against a vertically sheared flow that decreases with height will lose energy to that flow, while lee waves propagating against a shear flow that increases with height will take energy from that flow. The former mechanism is an important energy sink for lee waves (Waterman et al., 2014, 2021; Kunze & Lien, 2019). Similarly, horizontal straining of waves by the mesoscale eddy field can lead to significant energy exchange, and eventual wave dissipation in certain cases (Buhler & McIntyre, 2005). Because the Southern Ocean exhibits a vigorous and deep-reaching mesoscale eddy field, it may be a global hotspot for wave–mean interactions. However, numerical modelling support for this hypothesis is limited and observational evidence is almost non-existent for all but a few possible interaction mechanisms (Cusack et al., 2020).

One key wave–mean interaction in the Southern Ocean is the phenomenon known as the “critical level” (or “inertial level”; e.g., Booker & Bretherton, 1967). A critical level is a height at which the internal wave phase speed equals the horizontal mean flow speed and will be encountered when flow-trapped (e.g., lee) waves propagate upwards through a mean flow that decreases in magnitude with height above the bottom (Fig. 28; which usually occurs when the wave reaches the boundary of an eddy or jet), if the waves have not already dissipated via other means nearer the bottom (e.g., Nikurashin et al., 2013). During the propagation towards critical levels, the waves’ vertical wavelength decreases while their shear increases until, close to the critical level, instabilities lead to dissipation of the wave and the deposition of the wave momentum. Critical levels have also been suggested as a mechanism for the observed enhancement of dissipation around the edges of mesoscale eddies in Drake Passage (Sheen et al., 2015), with the potential for the wave momentum associated with tidally-generated internal waves to ‘spin up’ the eddies due to concomitant preferential dissipation of waves propagating in the direction of the mesoscale flow (Shakespeare & Hogg, 2019). Cusack et al. (2020) found significant energy transfers from internal waves propagating through eddy shear at a Drake Passage mooring, suggestive of a critical level type mechanism.

Many observational studies of wave–mean interactions in the Southern Ocean have been focused in regions of standing meanders downstream of major topographic obstacles (such as Kerguelen Plateau) that generate a vigorous eddy field (Sheen et al., 2015; Meyer, Polzin, et al., 2015; Waterman et al., 2021; Cyriac et al., 2023) because these are hotspots for key physical processes central to Southern Ocean dynamics (cf. § 4.1). Flow–topography interactions are elevated in these regions where the energetic jets of the Antarctic Circumpolar Current merge and split (Rintoul, 2018). In addition, the wind-energy input into near-inertial motions is high in these regions (§ 5.3). Thus, standing meanders are expected to be Southern Ocean mixing hotspots owing to the rich internal wave field generated from strong wind forcing and flow–topography interactions.

The elevated shear, strain and vorticity in the background flow in meanders are important factors in the evolution of internal waves. A timescale characterization of the various
processes expected to drive wave evolution suggests that the timescales associated with background flow advection and wave-mean flow interactions dominate dissipation timescales in the evolution of waves (Meyer, Polzin, et al., 2015; Waterman et al., 2021; Cyriac et al., 2023). This timescale analysis implies that some internal waves contribute to local mixing by dissipating locally, while most of the waves are advected away by the mesoscale flow and lead to dissipation downstream of the meander, in agreement with modelling studies (e.g., Zheng & Nikurashin, 2019) and theoretical descriptions (Shakespeare et al., 2021; Baker & Mashayek, 2021). The mixing driven by this far-field dissipation of internal waves has significant implications for the Southern Ocean stratification and watermass transformation (Meyer, Polzin, et al., 2015). Other potential mechanisms of wave–mean interactions in meander regions are the wave-capture (Meyer, Polzin, et al., 2015; Waterman et al., 2021) and near-inertial wave trapping (Meyer, Polzin, et al., 2015; Rama et al., 2022; Cyriac et al., 2023). Whether internal waves are located inside or outside fronts, jets and eddies controls which of these wave–mean interaction mechanism dominates.

5.3.3 High-latitude wave dynamics

In the Southern Ocean, the tidal frequency is everywhere less than the inertial frequency for the diurnal tide, and in the Ross and Weddell Seas (poleward of 74.5°S) for the most energetic semidiurnal tide, M2. In these regimes, internal tides are not freely propagating, but are instead generated as waves that are trapped near the bottom topography, either in the open ocean (bottom trapped waves; Rhines, 1970; Falahat & Nycander, 2015), or along the shelf (coastal trapped waves; Huthnance, 1978; Mysak, 1980). Coastal trapped waves can also be initiated by wind stresses and dense water outflows that produce sub-inertial oscillations (J. Adams & Buchwald, 1969; Marques et al., 2014; Liao & Wang, 2018). Unlike freely propagating waves that can travel across continental shelves and oceans, coastal-trapped waves must dissipate their energy near the shelf and slope and are thus a potential source of regionally important shelf mixing and mass transport (Musgrave et al., 2017). Trapped waves may also play an important role in modifying the amplitude of the surface tide in the Southern Ocean (§5.2.4). Coastal trapped waves propagate with the coast on their left in the Southern Hemisphere, with a form that is highly dependent on the characteristics of the topography and stratification (Schlosser, Jones, Musgrave, et al., 2019; C. W. Hughes et al., 2019).

Three general categories of coastal-trapped waves have been identified as important to Southern Ocean dynamics. In some regions, notably the Ross and Weddell Sea shelf breaks, the strongest currents are associated with coastal trapped waves forced by the diurnal tide (§5.2.1; J. H. Middleton et al., 1987; Whitworth & Orsi, 2006; Padman et al., 2009; Semper & Darelius, 2017). Coastal trapped waves of subtidal frequency have also been observed along shelf breaks (e.g., J. H. Middleton et al., 1982). Models suggest that outflows of Dense Shelf Water can excite these waves along the Antarctic continental slope (Marques et al., 2014). A third source for coastal trapped waves is associated with the co-location of critical slope (slope of the topography that matches the wave ray angle, and at which the generation of internal waves is most efficient; e.g., Becker & Sandwell, 2008) and critical latitude for the M2-semidiurnal internal tidal waves along the southern Weddell Sea shelf break (Robertson, 2001; Daac et al., 2009). Numerical modelling demonstrates that coastal trapped semidiurnal waves are generated in that region, leading to enhanced near-bed velocities at the shelf edge and thick bottom layers (∼100 m).

Coastal trapped waves are expected to affect mixing, cross-slope exchanges, ice shelf cavities, melt rates and sea ice concentration. Eddy diffusivities from both finescale (Daac et al., 2009) and microstructure (Fer et al., 2016) observations show elevated near bottom values at a southern Weddell Sea shelf-break location, attributed to the semidiurnal coastal trapped waves. Based on modelling, Marques et al. (2014) proposed that coastal trapped waves forced by dense-water outflows would affect benthic mixing and cross-slope water mass exchanges in the vicinity of sources of dense water outflows in the Weddell and Ross Seas.
Each of these processes depends on stratification and mean flow along the continental slope. Therefore, we expect seasonal modulation of the coastal trapped waves, which has been observed for coastal trapped waves forced by the diurnal tides (J. H. Middleton et al., 1987; Semper & Darelius, 2017). Sun et al. (2019) also observe the modulation of the seasonal ice shelf basal melt rate variability near the grounding zone by coastal trapped waves in East Antarctica. There is substantial potential for feedbacks between coastal trapped waves and background stratification and mean flow through associated mixing (§ 4.3) and tidal rectification (§ 5.2.3).
6 Climate trends and future projections

The Southern Ocean dynamic system is changing in response to global warming (§1) and changes in its atmospheric drivers. In recent decades, surface wind speeds have increased over the Southern Ocean (Young et al., 2011; Young & Ribal, 2019), together with a marked change in the Southern Annular Mode (Arblaster & Meehl, 2006; J. Toggweiler, 2009; D. W. Thompson et al., 2011), precipitation has decreased at lower latitudes and increased at higher latitudes (Manton et al., 2020), and evaporation has decreased over the Southern Ocean (Boisvert et al., 2020). This section reviews the key trends in the different components of the Southern Ocean dynamical system and projections for future trends where available.

6.1 Large-scale circulation

Long-term trends in large-scale Southern Ocean circulation arise directly from local changes in momentum and buoyancy fluxes at the surface, changes in freshwater input at the Antarctic margin, sea level rise, and modified inflows and outflows from the Atlantic, Indian and Pacific basins to the north. In addition, the complex interplay of large-scale circulation with the cryospheric, turbulent and wave phenomena provide important feedbacks that influence these trends.

Perhaps the most well-observed trend is the significant warming and freshening of the upper Southern Ocean (the top \(\sim 2000 \text{ m}\)), within and to the north of the Antarctic Circumpolar Current (Gille, 2008, 2014; Rintoul, 2018; Roemmich et al., 2015). To the south of the Antarctic Circumpolar Current, there is a mix of weak warming in West Antarctica and weak cooling trends elsewhere (Armour et al., 2016). The relative warming and freshening of the near-surface ocean increases both the upper ocean stratification and the meridional temperature contrast across the Antarctic Circumpolar Current, which has important consequences for mesoscale turbulence, convection and internal waves (discussed below). In terms of the Antarctic Circumpolar Current itself, thermal wind balance (Hogg, 2010) and modelling (Shi et al., 2020) suggests that its baroclinic volume flux should increase in response to enhanced meridional buoyancy gradients. Argo and satellite observations have shown an acceleration of the zonal flow on the northern edge of the Antarctic Circumpolar Current (Shi et al., 2021), yet it remains unclear whether this represents a strengthening of the Antarctic Circumpolar Current or just a southward shift of the subtropical gyres (A. L. Stewart, 2021). The position of the Antarctic Circumpolar Current appears to be stable, despite shifting westerly winds (Chapman, 2017). To the south of the Antarctic Circumpolar Current, enhanced heat and freshwater fluxes from the warming atmosphere and accelerating glacial melt — rather than significantly modifying the surface ocean — are being taken up by the deep ocean through modification of the properties of the deep waters formed in this region. The reduced density of shelf waters would be expected to lead to a weakening of the dense overflows and Antarctic Bottom Water formation (Silvano et al., 2018; Lago & England, 2019; Q. Li et al., 2023), and a warming and freshening of the Antarctic Bottom Water. Such changes to the Antarctic Bottom Water properties have indeed been observed (Purkey et al., 2019) along with an associated reduction in abyssal stratification (H. J. Zhang et al., 2021). However, changes to the northward volume flux of Antarctic Bottom Water (i.e., the abyssal overturning circulation) are presently not able to be measured with sufficient precision to detect climate trends, and are complicated by the influence of winds (A. L. Stewart et al., 2021).

As noted above, there is a significant trend in Southern Ocean surface winds, with the dominant mode of atmospheric variability over the Southern Ocean (the Southern Annular Mode) strengthening and contracting poleward. These increasing winds, despite a degree of eddy compensation (§2.4), are predicted to enhance the upper overturning circulation according to theory and numerical models (Meredith et al., 2012; Morrison & Hogg, 2013) but there is not yet observational evidence for such a change. In fact, recent inverse models based on tracer observations suggest that the upper overturning is currently weakening, fol-
lowing a period of strengthening in the 1990s (DeVries et al., 2017; Rintoul, 2018), but these changes may be due to natural variability rather than an anthropogenic climate change signal (H. Thomas et al., 2008). In terms of the abyssal branch of the overturning, as discussed in § 2.5, the predicted influence of increasing westerlies is uncertain as it depends on the balance of two competing influences. On the one hand, the wind-driven enhancement of eddies (§ 6.3) is expected to increase internal lee wave generation in the Southern Ocean (§ 6.4) and, thus, the deep ocean mixing and concomitant upwelling of Antarctic Bottom Water (D. P. Marshall & Naveira Garabato, 2008). On the other hand, the enhanced westerly winds will drive increased northward fluxes of upwelling mid-depth water in the Southern Ocean (expected to enhance the upper overturning, as described above) but diminishing the amount transported southward to feed Dense Shelf Water and Antarctic Bottom Water formation (Ito & Marshall, 2008; Nikurashin & Vallis, 2011; Shakespeare & Hogg, 2012). The projected weakening of the polar easterly winds (Neme et al., 2022) will also contribute to reducing Dense Shelf Water formation, due to the reduced northward export of sea ice away from Antarctica and subsequent build up of sea ice over the dense water formation sites (Timmermann et al., 2002; M. S. Dinniman et al., 2018). The relative influence of these different effects is challenging to assess even with state-of-the-art high-resolution global ocean models (and impossible with contemporary climate models) since they must be able to accurately represent both Antarctic Bottom Water formation, its northward isopycnal volume flux, and the internal waves driving mixing on ~ 1 km scales (§ 7.2; Trossman et al., 2016; Kiss et al., 2020; Yang et al., 2021). However, it is likely that the impact of increased Antarctic meltwater on the abyssal overturning (§ 6.3) will dominate over any wind-driven changes (Q. Li et al., 2023).

An additional uncertainty is the role of the Weddell and Ross Gyres in regulating Antarctic Bottom Water properties and transport (§ 2.3), and how this may change into the future. The classical view of gyres as entirely wind-driven would suggest that the gyres amplify in response to enhanced Southern Ocean winds (Armitage et al., 2018; Vernet et al., 2019; Auger, Prandi, & Sallée, 2022). However, this view ignores the role of changes in sea ice cover (§ 6.2), which mediates the transmission of wind stress to the ocean (Neme et al., 2021), and the role of buoyancy forcing in driving the gyre circulation (Hogg & Gayen, 2020). In situ measurements of the gyres are summer-biased and remote sensing similarly limited by sea ice cover, meaning there are insufficient observations of this region to diagnose time variability or long-term trends (Neme et al., 2021). Likewise, current-generation climate models have insufficient resolution to represent key dynamics, and projections of changes in the subpolar gyres vary significantly in both sign and magnitude between models (Q. Wang et al., 2013). As such, it is currently not possible to make a definitive statement as to how the subpolar gyres are likely to change in the near future.

A similar story applies for the Antarctic Slope Current. As noted in the review of A. F. Thompson et al. (2018), ‘our understanding of Antarctic Slope Current variability and how it responds to a changing climate is in its infancy’. The variability of the Antarctic Slope Current is influenced by a complex interplay of changes in wind stress, sea ice cover, tides and freshwater input (Moffat et al., 2008; Hazel & Stewart, 2019; A. L. Stewart et al., 2019; A. F. Thompson et al., 2020; Si et al., 2021), and the relative balance of these effects in different regions may lead to regionally-varying trends in the Antarctic Slope Current. Modelling studies suggest that increased meltwater input will result in a near-circumpolar strengthening of the Antarctic Slope Current, due to the increased density gradient between freshening shelf waters and salty offshore waters (Moorman et al., 2020; Beadling et al., 2022).

6.2 Cryosphere

The accelerating trend for reduction and loss of Antarctic ice shelves (negative mass balance or volume losses; Paolo et al., 2015) has the potential to influence future trends in global sea level rise, as weakening of the ice shelf “buttressing effect” (the resistance ice...
shelves provide to glacial outflow; Gudmundsson, 2013) is the dominant driver of Antarctic Ice Sheet mass loss via “dynamic thinning” (i.e., glacier flow acceleration; Meredith et al., 2019). The overall rate of Antarctic ice shelf loss is almost equally split between ocean-driven basal ice shelf melting and calving of icebergs from ice shelf fronts (Rignot et al., 2013; Depoorter et al., 2013; Greene et al., 2022). The most recent assessment uses satellite measurements to estimate a net mass loss from Antarctic ice shelves loss over 1997–2021 of \( \approx 5,900 \text{ Gt} \) due to calving and \( \approx 6,100 \text{ Gt} \) due to thinning (Greene et al., 2022). Over the past one to two decades, Antarctic ice shelves have gained net area at a rate of \( \approx 350–500 \text{ km}^2 \text{ yr}^{-1} \) (Greene et al., 2022; Andreasen et al., 2022), but this follows a sharp drop in area due to major calving events (e.g., of the Ross Ice Shelf in 2002; Lazzara et al., 1999), and the area in unlikely to recover before the next major calving events (estimated by the end of the 2030s; Greene et al., 2022). Basal meltwater fluxes over the past three decades have been \( \approx 15\% \) greater than the “steady-state” (i.e., the melt rate required to maintain the ice shelf in a state of equilibrium) and up to \( 40\% \) greater in the peak year of 2009 (Adusumilli et al., 2020).

There are strong regional variations in ice shelf trends. The mass balance of small- to medium-size, warm-water cavities fringing West Antarctica and certain parts of East Antarctica, such as Getz, Totten and Pine Island, are dominated by basal mass loss (Depoorter et al., 2013; Rignot et al., 2013), such that they produce a substantial proportion of net ice-shelf basal meltwater despite only occupying a relatively small fraction of the total ice-sheet area (Rignot et al., 2013; Adusumilli et al., 2020). In contrast, giant, cold-cavity ice shelves, such as the Ross and Filchner-Ronne, are dominated by the cycle of ice-front advance and calving, with high basal melt rates confined to the ice fronts and grounding lines (Rignot et al., 2013). Overall, shelf-front processes are the strongest drivers of mass balance for most ice shelves (Depoorter et al., 2013; Greene et al., 2022), although thinning has had a greater impact on the buttressing effect (Greene et al., 2022). Based on current trends, certain ice shelves will lose substantial proportions of their volumes by the end of the twenty-first century (Paolo et al., 2015). Under high emissions pathways for future warming (RCP8.5), greatly enhanced ice shelf surface melt is predicted, such that several ice shelves will experience surface melt intensities comparable or greater than those experienced by Antarctic Peninsula ice shelves prior to disintegration (Trusel et al., 2015; de Conto et al., 2021). This may be exacerbated by loss or reduction of a sea ice barrier to the open ocean, which is already a trend for West Antarctic ice shelves (Reid & Massom, 2022; Teder et al., 2022). However, there are large uncertainties in model projections of ice shelf loss relating to feedbacks initiated by warming temperatures and, hence, low confidence in the future ice shelf trends (Fox-Kemper et al., 2021).

Satellite records of Antarctic sea ice extent date back to 1979 (Parkinson, 2019). Despite the warming atmosphere, the annual maximum Antarctic sea ice extent had a positive, albeit weak, trend of 13,800 km\(^2\) yr\(^{-1}\) until the mid 2010s (J. Liu et al., 2023). A record maximum of 20.11 million km\(^2\) was reached in September 2014 (NISDC, 2023). The phenomenon of increasing Antarctic winter sea-ice extent during an epoch of global warming is known as the ‘Antarctic paradox’ (J. King, 2014). It has yet to be decisively explained. Studies have pointed to the unconstrained geographical setting of the Southern Ocean, allowing largely unhampered sea ice drift, the warming sea surface temperatures (X. Li et al., 2014), the increased input of cold freshwater near the ocean surface from increased basal ice-sheet melt (Bintanja et al., 2013), and the changes in local winds (P. R. Holland & Kwok, 2012), as the drivers. The forcings were largely attributed to large-scale modes forced by teleconnections from lower latitudes, including the Southern Annular Mode and El Niño-Southern Oscillation (e.g., Stammerjohn et al., 2008), whereas the cause of the large-scale trends remains a point of discussion. Natural variability together with anthropogenic forcing are key contenders (D. W. Thompson et al., 2011).

Dramatic Antarctic sea ice losses during both winters and summers came shortly after the 2014 record maximum sea ice extent (J. Turner et al., 2022). The losses culminated
in a record low of 18.4 million km$^2$ in annual maximum sea ice extent on 31st August 2016 (NISDC, 2023), and several consecutive records of minimum ice extent in following years, including the lowest ever recorded Antarctic sea ice extent of 1.8 million km$^2$ on 21st February 2023 (NISDC, 2023). These recent extremes in Antarctic sea ice match a significant increase in variability from about 2007 onwards, linked to changes in the balance of sea ice trends across different Antarctic regions (Hobbs et al., 2023). A number of studies are currently underway to assess the attribution of atmospheric versus oceanic forcing in driving the record minima (summer 2016–2017, February 2022 and February 2023; Schroeter et al., 2023; L. Zhang et al., 2022). Due to the extreme lows in Antarctic sea ice cover in recent years, there is currently no statistically significant net long-term trend in Antarctic sea-ice extent (Fogt et al., 2022; J. Liu et al., 2023).

6.3 Turbulence

Since the early 1990s there has been an increase in the mesoscale turbulence field in the Southern Ocean (Hogg et al., 2015; Martínez-Moreno et al., 2021). This has largely been attributed to strengthening westerly winds (N. C. Swart et al., 2015). There is evidence of time lag of two to three years between changes in the wind forcing and the response of eddy kinetic energy (Meredith & Hogg, 2006b; Screen et al., 2009; Morrow et al., 2010). As described in § 2.1, the time-mean flow of the Antarctic Circumpolar Current is at most weakly sensitive to the changes in wind stress, with the wind instead acting to energise the mesoscale turbulence field, which is a phenomena known as “eddy saturation” (e.g., Munday et al., 2013; Constantinou & Hogg, 2019). In this way, the mesoscale turbulence field modulates the Southern Ocean response to strengthening winds. A more regional view of the mesoscale turbulence field shows evidence for local variability, with hotspots of increased eddy kinetic energy in regions with topographic features (A. F. Thompson & Naveira Garabato, 2014), for example downstream of the Kerguelen Plateau (Rosso et al., 2015). Recent work using satellite altimetry has highlighted that these eddy hotspot regions in the Southern Ocean are strengthening in eddy kinetic energy on the order of 5% per decade (Martínez-Moreno et al., 2021). These eddy hotspot regions are often crucial for the uptake of heat and carbon (Langlais et al., 2017) and, hence, the trends in these regions will influence future climate. However, such eddy hotspots are not represented in current climate models.

Disentangling the process or processes driving trends in eddy kinetic energy is challenging. There is a large-scale warming trend in the most strongly eddying regions in the vicinity of the circumpolar current (§ 6.1). The local gradients in sea surface temperature are increasing (on average), which is associated with intensifying eddy activity. However, the strengthening of mean winds, which leads to outcropping and tilting of isopyncals and increased baroclinic and barotropic instability, also feeds the eddy kinetic energy (Martínez-Moreno et al., 2021). Further, in addition to forcing, mesoscale turbulence trends are affected by other Southern Ocean dynamical processes. For example, the observed mixed layer deepening and stratification intensification may directly influence the mesoscale turbulence field by varying the degree of baroclinic instability or the scale of the eddies produced (by changing the Rossby radius of deformation). Changes in the stratification may also indirectly affect mesoscale turbulence via other processes such as modulating internal wave drag (§ 6.4), or influencing sea ice formation and the production of deep water masses. Cryospheric trends will also affect the mesoscale turbulence field. For example, increasing ice shelf melt rates will lead to increasing stratification, a transient increase in sea ice area and subsurface warming (Bronselaer et al., 2018; Haumann et al., 2020; Q. Li et al., 2023), which will impact the mesoscale turbulence field.

Convection is expected to be strongly influenced by buoyancy forcing trends. The deepening of the mixed layer and strengthening of the underlying stratification may already be an indication of enhanced convective processes in this region. Recent studies also suggest that the mixed layer is becoming fresher due to global warming, driven by changes in the
precipitation-evaporation balance, accelerated melting and calving of Antarctic glaciers, and a more positive phase of the Southern Annular Mode (J. Zhang, 2007; de Lavergne et al., 2014). Freshening of the surface ocean around Antarctica will stabilise the water column, reducing the ability of the mixed layer to overturn and entrain underlying water, and making coastal and open ocean convection events less frequent (de Lavergne et al., 2014; Moorman et al., 2020).

There is growing evidence that cyrospheric trends have a significant impact on both open ocean and coastal convection. Observations show that the calving of a large iceberg reduced the rate of sea ice production in a coastal polynya by blocking the flow of sea ice (Snow et al., 2018). This change in surface buoyancy conditions reduced convection and Dense Shelf Water production, which subsequently reduced the density and volume of the local Antarctic Bottom Water. Another cryospheric effect is the outflow of meltwater from neighbouring ice shelves into coastal polynya regions. It has been observed that this can reduce nearby convection and the rate of Antarctic Bottom Water formation (Silvano et al., 2018). If the trend of ice shelves is towards more calving and melting, then we might expect less convection and dense water formation on the Antarctic margins (Q. Li et al., 2023). However, other forcing changes, such as the strengthening Southern Annular Mode, may be responsible for opening up other polynyas and open ocean convection regions near Maud Rise (Jena et al., 2019; Kurtakoti et al., 2018). Therefore, it is challenging to predict the response of convection to climate trends.

The trends in mixing are difficult and, in many cases, near impossible to assess. Issues with measuring mixing (§4.3) impede the direct tracking of mixing trends. However, some work can be done with identifying trends in sources of mixing. In the upper ocean, changes in wind stress and surface buoyancy forcing will likely induce modifications in the mixing. Indeed, changes in the mixed layer depth and stratification are already being noted, which indicate that mixing is already adjusting in these regions. Another example is that increasing wind stress can produce stronger Langmuir circulation and hence more mixing in the upper ocean. Trends in melt rates under ice shelves will change the amount of mixing occurring across the meltwater plume. In the interior ocean, trends in internal waves and stratification are hypothesised to modify the mixing rates. In the deep ocean, trends in the abyssal water mass properties and stratification will influence the buoyancy transport in the bottom boundary layer and associated mixing. It is extremely difficult to determine even the direction of these trends, given the various competing influences.

### 6.4 Gravity waves

The trend of increasing wind speeds and storminess over the Southern Ocean is influencing both the surface and internal wave climates. The enhanced winds are expected to lead to an increase in generation of near-inertial internal waves along storm tracks, with energy fluxes increasing proportional to wind stresses (∼1% per decade; Cuypers et al., 2013; Rimac et al., 2013). However, given the paucity of internal wave observations and the inability of current climate models to resolve internal waves, this predicted change has neither been directly observed nor modelled. By contrast, satellite observations show that surface wave amplitudes in the Southern Ocean are growing faster than in any other region (Hemer, 2010; Young et al., 2011; Young & Ribal, 2019; Meucci, Young, Aarnes, & Breivik, 2020; Timmermans et al., 2020). Over the satellite era, the Southern Ocean has regions in which the mean significant wave height has a positive trend of up to 10 mm per year and in most regions extreme waves (defined as waves with heights above the 90th percentile) are also increasing at up to 10 mm per year (over 1985–2018; Young & Ribal, 2019). Twentieth-century climate model ensembles give century-long trends (1901–2010) of 10–20 mm per decade in mean significant wave height (Meucci, Young, Aarnes, & Breivik, 2020; Meucci et al., 2023). Under the RCP8.5 high-emission scenario (Van Vuuren et al., 2011), the largest ensembles to date predict that by the end of the century there will be up to 15% increases in significant wave heights (Morim et al., 2019), 5-10% in low-frequency extreme wave events.
1 in 100-year significant wave height return period, i.e., waves with a 1% probability of occurring in a given year; Meucci, Young, Hemer, et al., 2020) and 50–100% in high-frequency events (return periods less than one year; Morim et al., 2021). These projected changes in the Southern Ocean surface wave climate extremes are consistent across different datasets and statistical approaches (Lobeto et al., 2021; O’Grady et al., 2021).

In addition to winds, changes to the ocean stratification will play a major role in modifying the future internal wave climate. Observations broadly show increasing stratification in the upper Southern Ocean and reducing stratification in the abyss, although these trends are highly variable (Armour et al., 2016; Yamaguchi & Suga, 2019; H. J. Zhang et al., 2021). Weakened abyssal stratification will tend to suppress the production of topographically-generated internal waves (i.e., internal tides and lee waves), for which the energy flux scales with the buoyancy frequency above the topography (Bell, 1975). In turn, reduced internal tide generation will tend to enhance the strength of the barotropic tide, since it dampens a key energy sink (§5.2.4). Weakened abyssal stratification also pushes generated waves closer to a linear regime where they can more readily radiate away from the topography without breaking (Legg, 2021). Thus, reduced stratification in the abyss would likely tend to reduce the mixing in this region. By contrast, the enhanced stratification in the upper ocean is likely to increase wave dissipation and mixing there, primarily because the waves vertical propagation is slowed and thus the waves have more time to break or otherwise lose their energy (O. M. Phillips, 1966; Kunze et al., 2006). The stratification changes will also cause significant variation in coastal trapped waves, which are a key component of tides along the Antarctic continental shelf (§5.3.3) and are known to be highly sensitive in both structure and amplitude to stratification (Semper & Darelius, 2017). For example, Skardhamar et al. (2015) model large changes in the energetics of diurnal coastal trapped waves due to seasonal changes in stratification along continental slopes (albeit in the North Atlantic). We expect similar responses at longer time scales in the Southern Ocean, as stratification evolves due to anthropogenic forcing. However, the exact response defies a simple description and will depend on the details of the flow and topography at a given location.

Trends in circulation and mesoscale turbulence will also influence the future internal wave climate in the Southern Ocean. For example, changes in the Antarctic Slope Current (§6.1) will modify the strength and structure of coastal trapped waves (Skardhamar et al., 2015; Semper & Darelius, 2017). Further, changes in bottom flow speeds due to mesoscale eddies (Martínez-Moreno et al., 2021) and (to a lesser extent) the Antarctic Circumpolar Current (Shi et al., 2021) are expected to enhance internal lee wave generation, albeit partially offset by the reduced stratification. Enhanced currents and eddies will also modify wave–mean interactions, wave dissipation and mixing in currently unknown ways.

The final factor influencing trends for both internal waves and tides is feedbacks from the cryosphere. Since ocean tides are resonant phenomena closely tied to the ocean basin geometry, tidal elevations and currents are sensitive to changes in ice shelf thickness and extent (Rosier et al., 2014). Changes to tides are largest near the locations where the ice shelves change but can also exhibit non-negligible far-field effects over time (Rosier et al., 2014; Padman et al., 2018). For example, the thinning or areal reduction of large ice shelves, such as the Ross and FIlchner-Ronne Ice Shelves, might alter tidal energy distribution, initially only close to the ice perturbation but then farther afield as the perturbation grows (Padman et al., 2018).

Our knowledge of changes to the internal wave climate under ice in the Southern Ocean is similarly poor. While there have been studies on the changing internal wave climate under sea ice in the Arctic (S. T. Cole et al., 2014, 2018; Fine & Cole, 2022), including long term changes linked to climate change, there are currently no similar records under Antarctic sea ice. In parts of the Arctic, decreasing sea ice cover and increasing open water conditions have been linked to increased internal wave energy (Martini et al., 2014; Dosser & Rainville, 2016). However, this increase in internal wave energy is partially offset by increased stratification and a shallower mixed layer (Guthrie et al., 2013), and by a decrease in sea ice roughness...
as the Arctic shifts from rougher multiyear ice to smoother first year ice that generates fewer internal waves and less internal wave energy (S. T. Cole et al., 2018). In the Southern Ocean, we can expect the internal wave field under sea ice to have high regional, seasonal and interannual variability, matching the increased variability in sea ice cover, but this is yet to be investigated (Hobbs et al., 2023).
7  Summary and future research priorities

In previous sections, we have reviewed current knowledge of the complex, dynamical system that is the Southern Ocean, organised around the four broad constituent parts: large-scale circulation, cryosphere, turbulence and gravity waves. Here, we draw together these components in a summary overview of the dynamical system (§7.1), before proceeding to frame the priorities for future research (§7.2).

7.1  Summary

At large scales, water masses circulate around the Southern Ocean in both the horizontal and vertical directions, driven by wind stresses and buoyancy fluxes at the ocean surface. In the horizontal plane, the eastward-flowing Antarctic Circumpolar Current (§2.1) wraps around the entire Southern Ocean. It is the largest ocean current in the world and formed from a network of mesoscale eddies, jets and fronts. Farther south, the westward-flowing Antarctic Slope Current (§2.2) extends around almost the entire coastline of Antarctica, sustained by buoyancy fluxes from the cryosphere. Between these two currents, the Weddell and Ross gyres (§2.3) are large-scale horizontal circulations, which play a role in exchanging cold waters around Antarctica with warmer waters in the Antarctic Circumpolar Current. In the vertical–latitudinal plane, water flows mostly along isopycnals in the meridional overturning circulation. In the upper cell of the overturning circulation (§2.4), upwelling Circumpolar Deep Water flows polewards at mid-depths and downwelling mode and intermediate waters flow equatorwards north of the Antarctic Circumpolar Current, which is controlled by large-scale surface forcing, and also by smaller-scale processes, such as eddies and isopycnal mixing. In the abyssal cell of the overturning circulation (§2.5), there is a southwards flow of lower Circumpolar Deep Water, which is transformed into Dense Shelf Water as it meets the Antarctic continental shelf under the influence of the cryosphere. The dense water sinks to become Antarctic Bottom Water, which flows northwards and upwells due to small-scale diapycnal mixing.

The Southern Ocean connects to the cryosphere through ice shelves (§3.1) and sea ice (§3.2). Ice shelves form floating lids on the ocean surface, enclosing cavities of ocean water that are sheltered from direct forcing by the atmosphere. The cavities interact with the large-scale Southern Ocean currents (§3.1.1), facilitated by mesoscale processes, such as eddies close to the ice shelf fronts (§3.1.2), buoyant plumes along the ice shelf base (§3.1.4) and tides (§3.1.3), and smaller scale effects, such as the finescale turbulence in the boundary layer at the ice shelf base and the influence of roughness in the ice base topography (§3.1.6). The varying nature of these interactions results in a rich spectrum of ocean cavity behaviours and associated ice shelf basal melt rates. The Southern Ocean contributes to iceberg calving by thinning and melting at the shelf front and base, and imposing forces on ice shelves, including flexure due to surface waves (§3.1.7). The resultant loss of mass to iceberg calvings is comparable to that lost through basal melting. It contributes to the freshwater flux into the Southern Ocean but its signature is delayed as it happens through the life cycle of icebergs ranging from rapidly melting bergy bits through to the massive tabular icebergs, which can persist for decades and travel great distances.

A large proportion of the Southern Ocean surface surrounding Antarctica is seasonally covered by sea ice. Much of the sea ice cover is dynamic, i.e., it drifts, and its drift motions are coupled to ocean surface currents (§3.2.1). Large-scale sea ice drift patterns mirror large-scale circulation, with predominantly eastward drift in the open ocean following the Antarctic Circumpolar Current and a relatively narrow band of westward drift close to the coast. The local composition of the sea ice cover affects the way it drifts in response to winds and currents. For instance, sea ice is most mobile and drift speeds are fastest in the marginal ice zone at the outskirts of the sea ice cover, where sea ice is relatively thin, ice floe concentrations are typically low, and, under the influence of surface waves, floe sizes small.
In contrast, sea ice resists motion and drift speeds are slowest close to the coast, where the ice is compact and floe–floe interactions create internal stresses.

Instabilities in large-scale Southern Ocean circulation create strong turbulence fields. At the mesoscale, these tend to form eddies, jets and fronts (§ 4.1), which are persistent, spatially localised features that interact with one another, and can be viewed as the building blocks of large-scale circulation patterns. They are driven by buoyancy forcing and wind stresses (§§ 4.1.1–4.1.2), which can act as a source or a sink, and are either exacerbated by sea ice drift or eliminated by stationary sea ice cover, e.g., fast ice. Mesoscale turbulence is coupled to stratification, which creates further linkages with the cryosphere through ice melt and altered surface heat fluxes in regions covered by sea ice. In general, mesoscale turbulence acts as the bridge between the global-scale circulation and small-scale turbulent processes in the Southern Ocean. It cascades energy back to the larger scale flows and down to smaller scale turbulence (§§ 4.1.4, 4.1.6).

At the submesoscale, turbulent convective events appear in plumes of enhanced vertical motion (§ 4.2). They can result from instabilities in the upper mixed layer (§ 4.3.1), which are strongest in winter due to buoyancy loss to the atmosphere in the sea ice free ocean and brine rejection during sea ice formation, and act to deepen the mixed layer during winter through small scale turbulent process. Polynyas are active areas for localised convective plumes (§§ 4.2.2–4.2.3). For instance, in certain coastal polynyas, where katabatics push newly formed sea ice northwards, plumes can stretch hundreds of metres to the ocean floor and become self sustaining by bringing warm Circumpolar Deep Water to the ocean surface within the polynya to determine dense water formation.

Waters of different properties are irreversibly mixed at small-spatial and short-temporal scales through three-dimensional turbulence (§ 4.3), which ultimately removes the energy that cascades downwards from larger-scale motions. Turbulent mixing occurs in the upper ocean surface boundary layer (§ 4.3.1), driven by surface waves, Langmuir circulation, resonant near-inertial waves, eddies, fronts and jets. Ice shelf fronts enhance mixing. Sea ice cover can suppress certain turbulent processes by reducing or removing surface waves and momentum fluxes, and enhance other processes by increasing density gradients. Diapycnal mixing is widespread throughout the Southern Ocean interior due to the downward energy cascade driven by internal waves (§ 4.3.2), and intensified at the ocean bottom (§ 4.3.3) by the local breaking of lee waves and internal tides near rough topography.

Gravity waves transport energy vast distances around the Southern Ocean. The Southern Ocean possesses the largest surface waves on the planet (§ 5.1.1), generated by storms over large fetches, and subsequently developing to swell as they travel across the ocean surface. They contribute to mixing of the ocean surface layer and influence air–sea momentum, heat and carbon fluxes (§ 5.1.2). They are influenced by eddies and jets (§ 5.1.3), which can strengthen or weaken the waves depending on their relative directions, and by sea ice (§ 5.1.4), typically in the marginal ice zone, which redistributes and dissipates wave energy, so that surface waves only reach tens to hundreds of kilometres into the sea-ice covered ocean. Southern Ocean tides (§ 5.2) and associated currents impact the cryosphere, e.g., basal ice shelf melting, and turbulent processes, e.g., as a source of ocean bottom diapycnal mixing. They also generate internal tides and waves. More generally, internal waves are generated by perturbations in the ocean density field (§ 5.3.1), typically at the ocean surface due to winds or the ocean bottom due to interactions between the mesoscale mean flow and topography. They carry energy and momentum into the ocean interior (§ 5.3.2), driving small-scale mixing and and sustaining large-scale overturning circulation circulation — thus closing the loop on the vast dynamic system that is the Southern Ocean.

### 7.2 Priorities

Many significant gaps exist in our understanding of the Southern Ocean. Of course, this is also true for the global oceans as a whole, but the gaps are (arguably) more significant for
the Southern Ocean because of its unique role and the challenges in observing its behaviour. As a result, we lack detailed understanding of many of the unique dynamical regimes of the Southern Ocean, including unlimited-fetch surface waves (§5.1), extensive sub-glacial ice cavities (§3.1.2), atypically small mesoscale eddies on the Antarctic continental shelf (§4.1), and extreme internal tide incoherence (§5.3.1). This lack of knowledge compromises our ability to model the Southern Ocean and, therefore, accurately predict future climate change and sea level rise. Such accurate prediction is the key to well-targeted mitigation and adaptation efforts that limit the environmental, social and economic impacts from climate change. Thus, improving our ability to model the Southern Ocean system and its response to anthropogenic forcing must is the overarching priority.

Indeed, over half the uncertainty in projections of global mean sea level is due to the melting of the Antarctic Ice Sheets (Kopp et al., 2014). Reducing this uncertainty requires advances on multiple fronts due to the range of processes that influence the melt rate (Cook, Nicholls, Vaňková, Thompson, & Galton-Fenzi, 2023). Necessary advances include predicting trends in the large-scale circulation and temperature of the Southern Ocean beyond the continental shelf, understanding the transport mechanisms that flux heat onto the shelf and into the ice cavities, and developing accurate parameterisations of the finescale convection and turbulence that ultimately melts the ice. Each of these components contributes to the overall uncertainty and needs to be addressed.

Similarly, a key contributor to uncertainty in global mean air temperatures on long timescales is the rate of heat storage in the abyssal ocean (Abraham et al., 2013). While most anthropogenic heat is currently stored in the upper ocean (Levitus et al., 2012), which overturns faster, the abyssal ocean is playing an increasing role and will be crucial to the long-timescale evolution of climate change. Heat storage in the abyssal ocean is in large part controlled by the properties and quantity of dense bottom waters formed in the Southern Ocean, which then flow north in the abyssal branch of the overturning circulation. As discussed in §6.1, it remains an open question whether this abyssal overturning will increase or decrease under climate change. The sign and magnitude of the trend is influenced by a host of processes including the poorly understood dynamics of the Ross and Weddell gyres, changes in sea ice cover and brine rejection, the small-scale convection that leads to dense shelf water, and the unknown distribution and magnitude of mixing in the abyssal Southern Ocean. Our lack of knowledge in all of these areas ultimately contributes to uncertainty in global climate change.

Addressing these substantive challenges requires a community effort that spans from developing theoretical understanding of individual processes, through to new observations and novel data analysis, innovative numerical methods, and finally putting these components together to further develop the global ocean–sea-ice models that are used in climate and sea level projections. We now describe the specific priorities in each of these areas that feed into addressing the uncertainties identified above.

7.2.1 Process-based models

Numerical and laboratory models of individual processes underlie much of our understanding of Southern Ocean physics. They facilitate the untangling of complex, coupled processes into their constituent parts and form the basis for mathematical theories and parameterisations that are then utilised in large-scale modelling and the interrogation of observational data. Often the processes of interest occur on scales that cannot be resolved in regional or global models.

In recent years, process-based numerical models have proved key in problems, such as quantifying the role of lee waves in Southern Ocean abyssal mixing (e.g., Nikurashin & Ferrari, 2011, 2013), the dynamics of eddy hotspots and upwelling (Barthel et al., 2022), polynya convection (Sohail et al., 2020), abyssal upwelling along topography (Drake et al., 2022), eddy saturation (Constantinou, 2018; Constantinou & Hogg, 2019), and Antarctic
Slope Current dynamics (Ong et al., 2023). Future priorities include investigating convection in ice shelf cavities, internal-wave–eddy interactions and mixing, and surface wave–sea ice interactions in the marginal ice zone. All of these processes currently lack a sufficiently complete theoretical description to permit their integration into large-scale models. In addition, while the inference of mixing made in the past from finescale parameterisations (§4.3) applied to observations is immensely valuable, key questions remain about the assumptions involved (Bluteau et al., 2013; Polzin, Naveira Garabato, Huussen, et al., 2014; Mashayek, Salehipour, et al., 2017; Gregg et al., 2018; Ijichi et al., 2020). These assumptions can be queried using idealised process-based models, and the resulting theory applied to improve the interpretation of extant and future observations.

Laboratory experiments have also played a major role in developing our understanding of key Southern Ocean processes, such as convection (e.g., Vreugdenhil et al., 2017; Gayen & Griffiths, 2022), wave breaking and air–sea exchange (e.g., Melville, 1996; Mayer et al., 2020), jet dynamics (e.g., Von Larcher & Williams, 2014; C. A. Smith et al., 2014), gravity currents (e.g., Griffiths, 1986), mixing and internal waves (e.g., Dossmann et al., 2016; Tan et al., 2022), and ice–ocean interactions (e.g., Aussillous et al., 2006; McCutchan & Johnson, 2022). Laboratory modelling has become less common in recent years, largely due to the relative cheapness and adaptability of numerical modelling. However, laboratory experiments remain a crucial tool in understanding many (especially multi-phase) systems where the governing dynamical or thermodynamical equations and/or boundary conditions are not necessarily known (e.g., sea ice, complex glacial topologies, sediment-laden plumes, air–sea gas exchange). In particular, experiments of melting ice faces (reviewed by McCutchan & Johnson, 2022) form the basis for our current glacial melt-rate parameterisations which are used to predict future sea level, but recent comparisons (Malyarenko et al., 2020; Rosevear, Galton-Fenzi, & Stevens, 2022) show that more studies are needed to examine different regimes (such as melting under the influence of tides; Richter et al., 2022). Other key next steps are the identification of thresholds between melting regimes and the development of parameterisations based on properties resolved in global models. Similarly, it is becoming vital that we better understand how the thermal and optical properties of sea ice (e.g., albedo, thermal conductivity, brine content; §3.2.3; Perovich, 1996; Light et al., 2003; Pringle et al., 2007) may change in the future as the climate warms, so that these effects can be included in global ocean–sea-ice models. As such, new facilities are being set up to study the thermodynamics of sea ice (e.g., M. Thomas et al., 2021; Hall et al., 2023) under carefully controlled laboratory conditions. Such investigations are likely to be critical in improving the accuracy of ocean-sea ice model projections of future climate scenarios and not simply tuned to describe present climate conditions.

7.2.2 Observations

The Southern Ocean, especially in the cryospheric zone, is one of the most poorly observed regions of the global ocean, due to the high cost and difficulty of accessing this region (Meredith et al., 2013). However, observations are crucial to our real-time monitoring of how the Southern Ocean system is changing. Observations also allow testing of process-based models and theories in realistic geophysical settings, and calibration of highly-parameterised regional and global numerical models. Without observations we would have no method of evaluating whether these numerical models (and their associated climate projections) are representative of the real world. Both in situ and satellite observations are crucial in performing these functions.

Satellite observations are exceptionally useful because they provide continuous observations in time with near complete spatial coverage, allowing measurement of, e.g., sea ice extent, ice sheet mass loss, surface wave fields, geostrophic eddies and currents, and ocean tides. New satellite missions now underway, e.g., Surface Water and Ocean Topography (SWOT; Morrow et al., 2019) and, Surface Wave Investigation Measurements (SWIM Aouf et al., 2020)), promise unprecedented resolution. This should improve our understanding
of the small eddies on the Antarctic continental shelf that are key to heat transport, and short-wave components of the surface wave field that are characteristic of the long-fetch conditions of the Southern Ocean. Such observations will be used directly in data assimilating models, and in the testing of theories and parameterisations for these small-scale processes.

In situ observations are vital to our understanding of processes occurring below the sea surface, which are invisible to satellites. Ship-based observation in the Southern Ocean is expensive and strongly biased towards more amenable summertime conditions, and easier to access regions (Newman et al., 2019). As such, we lack sufficient observations in many environments, such as beneath sea ice cover and in ice shelf cavities, and during rough weather conditions, which are crucial for determining ocean mixing, ice sheet melt rates and dense water formation. However, new platforms are coming online that are starting to fill some of these gaps. For example, deep Argo and other floats are now available that profile year-round and under sea ice, which should greatly expand data coverage in the far Southern Ocean (Johnson et al., 2022; van Wijk et al., 2022). In addition, creative solutions such as animal-borne data acquisition are becoming more widespread (Roquet et al., 2014; Foppert et al., 2019). Through-ice moorings are also providing valuable insights into hydrography, currents and turbulence (e.g., Arzeno et al., 2014; Davis & Nicholls, 2019; Stevens et al., 2020; Hattermann et al., 2021; Herráiz-Borreguero et al., 2013), and some measurements are now being obtained by autonomous underwater vehicles including submarines and gliders (e.g., Nicholls et al., 2006; Gwyther et al., 2020; Schmidt et al., 2023b). In addition, advances in surface radar enable highly resolved (in space and time) measurements of the ice shelf base (Vančová et al., 2021) that are sufficiently accurate to identify tidal modulation of melt rates (Sun et al., 2019). In terms of ocean mixing, the development of microstructure profiling Argo floats (Roemmich et al., 2019) and gliders (Wolk et al., 2009) is a particularly enticing possibility. Current and future trends in mixing intensity, potentially associated with trends in winds and eddy kinetic energy (e.g., Sheen et al., 2014; Whalen et al., 2018; Martínez-Moreno et al., 2021), remain open questions, which more observations with such platforms can help to constrain. However, there remains an urgent need to prioritise longer term continuous and sustained in situ measurements to permit the detection and analysis of long-term trends and seasonal variability. For example, a Southern Ocean equivalent of the North Atlantic RAPID array (Cunningham et al., 2007) to directly monitor the large-scale circulation.

While there will always be demand for new and improved observations, it is equally important that we make better use of the data we already have, both in terms of science (e.g., developing novel analysis methodologies) and data management. On the science side, efforts are underway to develop novel methods of extracting Southern Ocean bottom pressures and abyssal circulation from gravimetric satellite observations (e.g., GRACE; Wouters et al., 2014), an approach which has proven successful in the North Atlantic Ocean (Landerer et al., 2015). Significant work is also being done to measure the Southern Ocean internal tide field and associated mixing from existing satellite altimeter data (Z. Zhao et al., 2018), including addressing the challenge of wave dephasing due to the strong Southern Ocean eddy field using machine learning methods (H. Wang et al., 2022; Egbert & Erofeeva, 2002). Such efforts will be crucial in addressing the observation deficit in the Southern Ocean.

In terms of data curation, it is essential that all data generated by the Southern Ocean community is managed in accordance with the FAIR data principle; that is, data should be findable, accessible, interoperable and reusable (Wilkinson et al., 2016). Genuine accordance with this principle is essential for the community to gain maximum benefit from new and existing Southern Ocean data, and ensure cost-effectiveness for funding agencies. Community data curation efforts such as the Southern Ocean Observing System (SOOS; Newman et al., 2019) and related projects, such as NECKLACE (collating ice shelf melt data; Cook, Nicholls, Vančová, Thompson, & Stewart, 2023) play a key role in this effort, and should be further expanded.
7.2.3 Regional and global models

Numerical ocean and climate models are our primary tool for future climate projection and operational ocean forecasting. These models are inevitably limited by their finite spatial resolution, with typical grid sizes of 1° in current generation global climate models (e.g., CMIP6; Roberts et al., 2020) and up to 1/12° in current global ocean-only models (e.g., Kiss et al., 2020) and ocean state estimates (e.g., Lellouche et al., 2018). Processes smaller than the grid scale must be parameterised in such models, i.e., a mathematical model for the process must be formulated, calibrated (e.g., with observations and process-based models) and implemented (H. T. Hewitt et al., 2020). As outlined in this article, many of these unresolved processes are crucial to the climate state (e.g., diapycnal mixing, deep convection, eddies) and yet many are still not fully understood. To some extent, these challenges are resolved by running ever-higher resolution models as computational power increases, avoiding the need for parameterisation. For example, ocean model grids finer than 1 km are needed to resolve eddies and their associated heat transport on the Antarctic continental shelf (Hallberg, 2013; A. L. Stewart & Thompson, 2015) and such resolutions are now feasible for very short global ocean simulations (Rocha et al., 2016; A. L. Stewart et al., 2018). Even if model speedups due to using graphical processing units (GPUs) render 1 km-resolution simulations close to ‘routine’ (e.g., Oceananigans.jl model; Ramadhan et al., 2020), processes such as diapycnal mixing will still not be resolved and there remains a need to parameterise other larger-scale processes for longer duration simulations.

As such, there is an urgent need for improved parameterisations of a number of key processes in large-scale ocean and climate models, including mixing (Melet et al., 2015), eddies (H. T. Hewitt et al., 2020), convection (Sohail et al., 2020), internal wave–eddy interactions and momentum transfer (Shakespeare & Hogg, 2019), surface wave–sea ice interactions (Bennetts, Bitz, et al., 2022a), and surface and bottom submesoscales (Gula et al., 2022). Of these priorities, the representation of diapycnal mixing is recognised as particularly vital as it controls the strength and variability of the overturning circulation realised in such models (Melet et al., 2015). While static maps of mixing have been developed (de Lavergne et al., 2020), and parameterisations of some specific mixing processes have been implemented in global models (e.g., ice waves; Stanley & Saenko, 2014), development of a dynamically evolving representation of diapycnal mixing is a key priority. In developing such parameterisations, particular care should be taken to account for the unique dynamics of the Southern Ocean (e.g., §5.3.3) that lead to different mixing properties. Due to the changing climate, it is also essential that any parameterisation is physically based, and includes all relevant coupling with other processes. For example, empirical parameterisations based on present-day observations may fail in future ocean states, which will exhibit different stratification, mean currents and basin geometry (due to ice shelf and sea level changes). Further, it is vital that all parameterisations are ‘scale-aware’, i.e., they adapt to the model resolution, so as to avoid both parameterising and resolving the same process, and also to avoid the parameterisation negatively impacting the resolved phenomena.

The lack of scale-awareness is a well-known problem with the widely used Gent and Mcwilliams (1990) mesoscale eddy parameterisation at intermediate “eddy-permitting” resolutions (Hallberg, 2013; Jansen et al., 2019). While largely abandoned at the highest model resolutions, mesoscale eddy parameterisation remains important for lower resolution ocean and climate models, and the Gent and Mcwilliams (1990) parameterisation is arguably insufficient (H. T. Hewitt et al., 2020). To address such challenges, there is a recent move towards machine learning approaches to parameterise eddies (Bolton & Zanna, 2019; Zanna & Bolton, 2020, 2021; C. Zhang et al., 2023) as an alternative to simple mathematical models. The concept of these approaches is for the algorithm to ‘learn’ the governing physics of mesoscale eddies from eddy-resolving ocean models, with the resulting formulae then applied in lower resolution ocean and climate models. Such novel methods (although not
Another approach for incorporating the effects of the mesoscale ocean eddies in climate projections is the ‘brute-force’ approach of actually resolving them. By harnessing the power of graphical processing units (GPUs), ocean models can now achieve speedups of order 50–100 times compared to state-of-the-art ocean models that run on parallel arrays of central processing units (CPUs) (see, e.g., Ramadhan et al., 2020). With such speedups, climate projections at eddy-resolving ocean resolutions (e.g., 1/12°) may soon become routine and, in this way, any sort of mesoscale parameterisation may be rendered redundant in the near future.

As noted above, the ocean state in large-scale models is highly sensitive to mixing. As a result, elimination of unintended and spurious ‘numerical mixing’ is of equal importance to the accurate representation of physical mixing. Numerical mixing occurs due to the discrete representation of smoothly varying tracers, such as temperature and salinity, which are mapped onto gridpoints at each model timestep. Discrete mapping causes an unintended redistribution of tracer between adjoining grid cells (i.e., mixing), e.g., as the water column sloshes up and down due to the passage of an eddy or wave (Petersen et al., 2015; A. H. Gibson et al., 2017; Megann et al., 2022). Numerical mixing is difficult to quantify in complex models, but assessments that do exist suggest it can be significant, including in the eddying regions of the Southern Ocean (Holmes et al., 2021). This problem is important for the correct representation of Antarctic Bottom Water and the abyssal overturning circulation in the Southern Ocean. The amount of numerical mixing is closely tied to the vertical coordinates used in large-scale models and significant resources at major modelling centres are being devoted to determining an optimal vertical coordinate (e.g., A. Gibson, 2019; Klingbeil et al., 2019; Griffies et al., 2020; Wise et al., 2022). It is anticipated that these efforts will lead to increased model accuracy without the significantly increased computational expense.

A further priority in large-scale modelling is the incorporation of additional missing components of the Earth system. This includes the incorporation of ice shelf cavities and iceberg melt into ocean models, and the coupling of ice-sheet models with their ocean–sea ice counterparts (e.g., Favier et al., 2019; Gladstone et al., 2021; Kreuzer et al., 2021). Both of these efforts are likely to prove crucial to the accurate projection of future melt rates, but come with substantial computational challenges (Mathiot et al., 2017). Another missing feature in most ocean and climate models is an explicit representation of the ocean tides (Richter et al., 2022). Explicit inclusion of tidal currents in models, including baroclinic tides, would improve representation of benthic, mid-water and ice-base mixing, and the generation of rectified flows that help ventilate cavities (Makinson & Nicholls, 1999). However, inclusion of tides in a global model is not as simple as turning on the gravitational forcing, since the amplitude of tides are set by a balance between the forcing and the drag at the seafloor, some of which occurs at unresolved scales (Arbic et al., 2018). Therefore, the inclusion of tides requires the further development and co-implementation of additional parameterisations, supported by observations and process-based models.

### 7.3 Closing remarks

In many respects, the Southern Ocean is the final frontier of ocean science. It is a vast, poorly observed, inhospitable and almost untouched region that has fascinated humankind since the discovery of Antarctica in the 1820s. Scientific interest in the Southern Ocean has grown exponentially in recent times, along with understanding of the control Southern Ocean dynamics exert on global climate and climate change. However, progress has sometimes been stymied by a lack of effective communication between scientists in different disciplines and using different methodologies. This holistic review of Southern Ocean dynamics has sought to provide a common language and knowledge-base across the Southern Ocean scientific community to facilitate future knowledge-sharing and collaboration. Such
collaboration is critical to address the key scientific priorities identified above that span the disciplines of mathematics, fluid mechanics, software engineering, glaciology and oceanography, and methodologies as diverse as laboratory experiments of individual processes through to numerical modelling of the entire Southern Ocean system. All of these disciplines and methodologies — and many more — have a crucial role to play in accelerating our understanding of Southern Ocean dynamics in the years ahead, and thereby improving our ability to predict ocean and climate change. This outcome is critical for the global community, and indeed forms one of the goals of the United Nations Decade of Ocean Science 2021–2030. Facilitated by this review, we encourage the entire Southern Ocean community to come together to support this objective.
Open Research

No new data was used in this article. All new figures appearing in this review were made using publicly available data as cited in the relevant caption.

Acknowledgments
We acknowledge the receipt of the Elizabeth and Frederick White Research Conference Award from the Australian Academy of Science, which funded the initial meeting that led to this review.

The Australian Research Council supported the authors via its various grant schemes (Grant numbers: FL150100090; DE170100184; DP190100494; DP190101173; FT180100037; FT190100404; FT190100413; DP200102828; LP200100406; SR200100008; DE21010004; DE220101027; DP230102994).

LGB acknowledges support from the Australian Antarctic Science Program (AAS4528). GJS acknowledges support from the Banting Postdoctoral Fellowship through funding (180031). FM acknowledges support by the Marsden Fund (grant numbers 18-UOO-216 and 20-UOO-173) administered by Royal Society Te Aparangi. CS, AM, and FM acknowledge the Ministry of Business, Innovation and Employment Antarctic Science platform (ANTA1801).

Figures were produced using computational resources and data hosted by the National Computational Infrastructure, Canberra, Australia, a facility of the Australian Commonwealth Government.

Stacey McCormack produced the schematics. Anton Steketee and Sean Chau contributed to Fig. 9. Alison Kohout provided data for Fig. ???. Claire Yung provided code used to produce Fig. 18.

In keeping with the spirit of this review, we also acknowledge the many collaborative research consortia that supported this work including the: Consortium for Ocean Sea Ice Modelling in Australia, ARC Centre of Excellence in Climate Extremes, ARC Centre for Excellence in Antarctic Science, Australian Antarctic Partnership Program, and the Australian Government’s National Environmental Science Program Climate Systems Hub.
References


Arzeno, I. B., Beardsley, R. C., Limeburner, R., Owens, B., Padman, L., Springer, S. R.,


Goff, J. A., & Arbic, B. K. (2010). Global prediction of abyssal hill roughness statistics for use in ocean models from digital maps of paleo-spreading rate, paleo-ridge orientation,


Jendersie, S., Williams, M. J., Langhorne, P. J., & Robertson, R. (2018). The density-driven


Karsten, R., Jones, H., & Marshall, J. (2002). The role of eddy transfer in setting the


Kreuzer, M., Reese, R., Huiskamp, W. N., Petri, S., Albrecht, T., Feulner, G., & Winkelmann, R. (2021). Coupling framework (1.0) for the pism (1.1. 4) ice sheet model and the mom5 (5.1. 0) ocean model via the pico ice shelf cavity model in an antarctic domain. Geosci. Model Dev., 14(6), 3697–3714.


Melsheimer, C., & Spreen, G. (2019). *AMSR2 ASI sea ice concentration data, Antarc-


Stokes, G. G. (1880). Considerations relative to the greatest height of oscillatory irrotational waves which can be propagated without change of form. *Mathematical and physical papers*, 1, 225-228.


Wei, W., Blankenship, D. D., Greenbaum, J. S., Gourmelen, N., Dow, C. F., Richter, T. G., ... others (2020). Getz ice shelf melt enhanced by freshwater discharge from beneath


